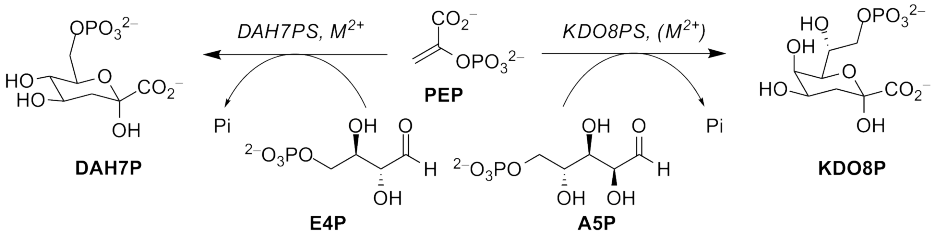
Substrate specificity and mutational studies of KDO8PS

A thesis submitted in partial fulfilment
of the requirements for the degree
of
Doctor of Philosophy in Biochemistry
in the Department of Chemistry
by Timothy Murray Allison

January 2012

Abstract

The enzyme 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS) catalyses the stereospecific aldol-like condensation between phosphoenolpyruvate (PEP) and the five-carbon sugar D-arabinose 5-phosphate (A5P). This is the first biosynthetic step in the formation of 3-deoxy-D-*manno*-octulosonate (KDO), an essential lipopolysaccharide component of all Gram-negative bacteria. KDO8PS is evolutionarily related to the shikimate pathway enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS), which catalyses a similar condensation reaction between PEP and the four-carbon sugar D-erythrose 4-phosphate (E4P), in the first step of the shikimate pathway to aromatic compounds in plants and microorganisms. As well as being a one-carbon shorter substrate, E4P has the opposite C2-OH configuration to A5P. While there are both metal-dependent and metal-independent forms of KDO8PS, in contrast, all DAH7PS are metal-dependent enzymes.



Little is understood about the key sequence features that distinguish KDO8PS and DAH7PS. These features, particularly those that contribute to A5P or E4P binding, are thought to be responsible for the differences in substrate specificity between the two enzymes. This thesis describes the

functional and structural studies of KDO8PS mutants to examine the roles of these residues, using the metal-dependent KDO8PS from *Acidithiobacillus ferrooxidans* and the metal-independent KDO8PS from *Neisseria meningitidis*.

In Chapter 2 an extensive KDO8PS and DAH7PS sequence analysis is presented. The results, which identify sequence conservation in both enzymes, are discussed in the context of the $(\beta/\alpha)_8$ TIM-barrel structure. Some of the differences in conservation between the two enzymes were highlighted as being obvious in having a role or contributing to the different substrate selection preferences of the two enzymes, such as an extended β7α7 loop in KDO8PS, and motif differences on the β2α2 and β4α4 loops. A similar analysis was also used to compare metal-dependent and metal-independent KDO8PSs, and it was found the two forms differ in the conservation of only three residues.

Chapter 3 describes the characterisation of *A. ferrooxidans* KDO8PS (*Afe*KDO8PS) and investigates aspects of metal dependency in KDO8PS. The enzyme was found to be metal dependent, and like all other KDO-8PS enzymes, to possess a tetrameric quaternary structure, and display tight substrate specificity. The β8α8 loop was found to have a critical role in binding and positioning the substrates, and *Afe*KDO8PS could not be engineered to be a metal-independent enzyme.

The role of the KDO8PS-conserved KANRS motif, present on the β2α2 loop and one of the main contributors to the A5P binding site, is probed in Chapter 4. Individual residues of the motif were mutated to investigate function, and the motif was converted to the equivalent motif found in DAH-7PS (KPRS). It was found that the Lys plays a critical role in enzymatic catalysis, and is likely intimately involved in the enzyme mechanism. The Asn residue of the motif in KDO8PS was found to be an important contributor to KDO8PS stereospecificity.

The work described in Chapter 5 investigates the role of the β7α7 loop in KDO8PS. This long active-site loop, which exists in a shorter version in DAH7PS, was found not to be essential for catalysis in KDO8PS, but was necessary for efficient catalysis. The two conserved residues on the loop

provide interactions to A5P, but the presence of the extended loop as a whole was found to be most important for catalytic efficiency.

In Chapter 6 a conserved residue on the *re* face of PEP is investigated. In KDO8PS the residue is conserved as Asp, and in DAH7PS the same residue is conserved as a Glu. Mutational analysis found that in KDO8PS the Asp residue appears to be important for enzyme activity but unimportant for PEP binding. Mutating this Asp in KDO8PS to Glu was accommodated by KDO8PS, but it was found its introduction could potentially be optimised by coupling the change with mutation to other conserved differences.

In KDO8PS, one of the interfaces between adjacent subunits in the tetrameric structure is partially composed of a conserved sequence motif, PAFLxR. In Chapter 7, the roles of the residues in this motif are explored. The Arg of the motif was found to be important for A5P binding. The equivalent (and also conserved) motif in DAH7PS is GARNxQ, and mutation of residues in the KDO8PS motif to the equivalent residues in DAH7PS was tolerated by KDO8PS, but negatively impacted upon the enzyme kinetic parameters. The sequence features investigated in the other chapters were combined with those to the subunit interface to create a DAH7PS-like protein. This extensively engineered protein lost all KDO8PS activity, but nor did it gain DAH7PS activity.

Lastly, in Chapter 8 the results from all chapters are reviewed and ideas are discussed for advancing the research presented in this thesis.

Acknowledgements

Many people need to be acknowledged, each for their own unique, direct or indirect, contributions to the work presented in this thesis.

Thank you to past and present members of the Parker research group for providing entertainment each day, and particularly to Dr Fiona Cochrane, Jeffrey Yeoman, Evan Nimmo and Ben Gloyne, whose work this thesis builds upon in many places. Despite your transient nature, thank you for your help and advice: Dr Sarah Lundy, Dr Steve McNabb, Dr Richard Hutton, and Dr Ali Reza Nazmi. To all of the inhabitants of room 858 over the years: thank you for the enlightening conversations and arguments, particularly Dr Aidan Harrison. Special thanks to Dr David Tran, Sebastian Reichau, Dr Scott Walker, Dr Daniel Packwood, Sarah Wilson-Coutts and Gerd Mittelstädt for helpful discussions on all manner of things. The gossip (and mass spectrometry services) provided by Dr Marie Squire will be missed.

Thank you to our cousins in Biology for always letting me in the lab door, and particularly to Dr Sean Devenish, Dr Andrew Muscroft-Taylor, and Dr Grant Pearce, for getting me started and always making available your expertise and equipment.

Thank you to the Biomolecular Interaction Centre, New Zealand Institute of Chemistry and New Zealand Society of Biochemistry and Molecular Biology for assistance by way of travel funding to attend conferences, the University of Canterbury College of Science for my doctoral scholarship, and the New Zealand Synchrotron Group for funding access to the Australian Synchrotron.

Thank you to my family for taking care of the domestic elements of life so I could concentrate on this.

Considerable thanks must go to Professor Geoffrey Jameson for his willingness to share his extensive chemical and crystallographic knowledge, for the exchange of ideas, and hospitality. Special thanks also to Dr Renwick Dobson, for introducing me to the Australian Synchrotron from which I have benefited greatly, and also for your help ("just chuck them in") in the early days. Particular thanks to Michael Hunter for being always willing, if not too keen, to exchange and argue ideas.

To the most hard-working supervisor in the world, Emily: you have taught me so much over the past five years about biochemistry, science, and many other subjects. This experience has shaped me and will stay with me for the rest of my life, for which I have you to thank.

But lastly, an enormous thank you to Penelope Cross. For all the special memories, gossip, the wedding, conference fun, many, many lab hours, and support and encouragement. It has been fantastic to have gone through the PhD experience with you.

Contents

Abbreviations

2dE4P 2-deoxyerythrose 4-phosphate.

2dR5P 2-deoxyribose 5-phosphate.

A5P D-arabinose 5-phosphate.

***Aae*KDO8PS** *Aquifex aeolicus* KDO8PS.

AEC anion-exchange chromatography.

***Afe*KDO8PS** *Acidithiobacillus ferrooxidans* KDO8PS.

***Ape*DAH7PS** *Aeropyrum pernix* DAH7PS.

***Apy*KDO8PS** *Aquifex pyrophilus* KDO8PS.

AUC analytical ultracentrifugation.

BTP 1,3-bis[tris(hydroxymethyl)methylamino]propane.

C3 Collaborative Crystallisation Centre.

CD circular dichroism.

CMP cytidine monophosphate.

CTP cytidine triphosphate.

DAH7PS 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase.

DNA deoxyribonucleic acid.

DSF differential scanning fluorimetry.

DTT dithiothreitol.

E4P D-erythrose 4-phosphate.

***Eco*KDO8PS** *Escherichia coli* KDO8PS.

EDTA ethylenediaminetetraacetic acid.

G6P glucose 6-phosphate.

HIC hydrophobic-interaction chromatography.

HMM hidden Markov model.

***Hpy*DAH7PS** *Helicobacter pylori* DAH7PS.

***Hpy*KDO8PS** *H. pylori* KDO8PS.

IPTG isopropyl β-D-1-thiogalactopyranoside.

ITC isothermal titration calorimetry.

KDO 3-deoxy-D-*manno*-octulosonate.

KDO8P 3-deoxy-D-*manno*-octulosonate 8-phosphate.

KDO8PS 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase.

LB lysogeny broth.

LPS lipopolysaccharide.

MSA multiple sequence alignment.

***Mtu*DAH7PS** *Mycobacterium tuberculosis* DAH7PS.

MWCO molecular weight cut-off.

***Ngo*KDO8PS** *Neisseria gonorrhoeae* KDO8PS.

***Nme*DAH7PS** *Neisseria meningitidis* DAH7PS.

***Nme*KDO8PS** *N. meningitidis* KDO8PS.

NMR nuclear magnetic resonance.

PCR polymerase chain reaction.

PDB Protein Data Bank.

PEP phosphoenolpyruvate.

***Pfu*DAH7PS** *Pyrococcus furiosus* DAH7PS.

R5P ribose 5-phosphate.

RMSD root-mean-square deviation.

Ru5P ribulose 5-phosphate.

SDS sodium dodecyl sulfate.

SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis.

SEC size-exclusion chromatography.

SOC super optimal broth.

TAE Tris-acetate-EDTA.

TIM triose phosphate isomerase.

***Tma*DAH7PS** *Thermotoga maritima* DAH7PS.

Index of figures

Index of tables

Publications

Parts of this thesis have been published in the following publications:

- **Allison, T. M.**, Yeoman, J. A., Hutton, R. D., Cochrane, F. C., Jameson, G. B., and Parker, E. J. (2010) Specificity and mutational analysis of the metal-dependent 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Acidithiobacillus ferrooxidans*. *Biochim. Biophys. Acta, Proteins Proteomics 1804*, 1526–1536.
- **Allison, T. M.**, Hutton, R. D., Cochrane, F. C., Yeoman, J. A., Jameson, G. B., and Parker, E. J. (2011) Targeting the role of a key conserved motif for substrate selection and catalysis by 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochemistry 50*, 3686–3695.
- **Allison, T. M.**, Hutton, R. D., Wanting, J., Gloyne, B. J., Nimmo, E. B., Jameson, G. B., and Parker, E. J. (2011) An extended β7α7 substrate-binding loop is essential for efficient catalysis by 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochemistry 50*, 9318–9327.

Chapter 1

Introduction

1.1 Gram-negative bacteria cell wall biosynthesis

1.1.1 The cell wall of Gram-negative bacteria

Gram-negative bacteria possess a cell wall that surrounds their inner cell membrane, which is composed of a peptidoglycan layer encased by an outer asymmetric lipid membrane.[1] The inner leaflet of the outer lipid membrane is composed of various glycerophospholipids, and the outer leaflet is composed of a lipid component decorated with outward-facing polysaccharides. The combined polysaccharides and lipid of the outer leaflet is called the lipopolysaccharide (LPS) or endotoxin, and is often associated with the pathogenicity of Gram-negative bacteria,[2,3] many species of which are notable for their medical relevance.

The chemical make-up of the endotoxin is often unique to species and strains, but some parts are conserved across nearly all Gram-negative bacteria. The endotoxin comprises three distinct regions: the O-antigen, core oligosaccharide and Lipid A (Figure 1.1).[4] The core oligosaccharide

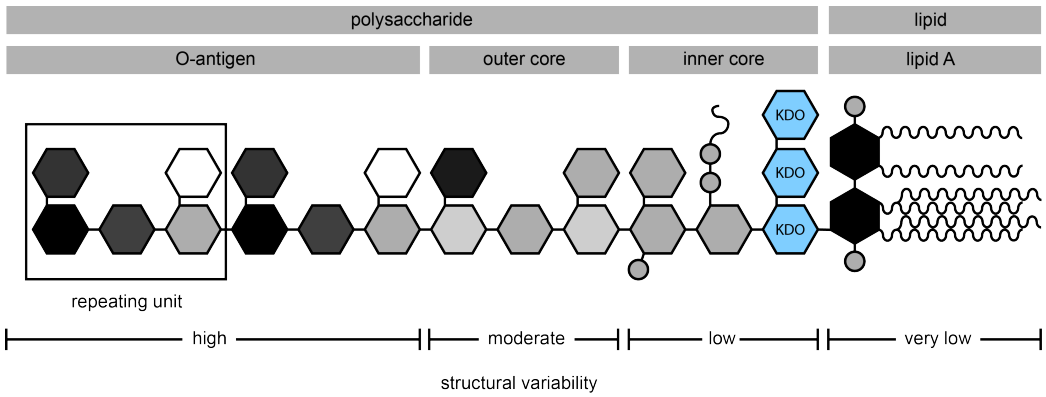


Figure 1.1: Schematic representation of the architecture of an enterobacterial wild-type lipopolysaccharide. Substituents may be present in non-stoichiometric amounts, e.g. the third KDO of the inner core region, fatty acids, and phosphoryl residues. The KDO molecules are coloured blue. Figure recreated from Wiese et al. (1999).[2]

region (composed itself of both an inner and outer region) is found between Lipid A and the structurally diverse O-antigen region. The O-antigen region determines the antigenic specificity of the strain.[4] Although in different species the core oligosaccharide is composed of a diverse number of different monosaccharides, the inner core immediately adjacent to Lipid A always contains one to three 3-deoxy-D-*manno*-octulosonate (KDO) molecules.[5,6]

The importance of the conserved KDO molecules in the inner region of the core oligosaccharide has been shown by studies that demonstrate, with few exceptions,[7,8] a minimal decoration of Lipid A with KDO is required for growth of Gram-negative bacteria.[4,9,10] This is because without KDO, the Lipid A precursor (IV_A) is not decorated with secondary acyl chains or glyco-additions, and the rate of transport of Lipid IV_A to the outer membrane is severely reduced.[7] The additional complex glycoforms of the outer core and O-antigen sugars are not needed for growth, but protect bacteria from antibiotics and are required for virulence.[9,11]

The dependence of cell viability and toxicity on KDO, and that the KDO biosynthetic pathway is found only in bacteria and plants,[i,14] makes KDO biosynthesis a potential target for antibacterial drug design.[6] Inhibition of this crucial biosynthetic pathway alone may be an efficacious strategy, but targeting it in combination with existing antibiotics may broaden the spectrum of activity of those with low outer-membrane permeability.

1.1.2 LPS-KDO biosynthetic pathway

The biosynthesis of KDO (Figure 1.2) begins with the isomerisation of ribulose 5-phosphate (Ru5P), an end product of the pentose phosphate pathway, and also an intermediate in the Calvin cycle. Ru5P, a ketose, is isomerised to the aldose D-arabinose 5-phosphate (A5P) by the enzyme A5P isomerase (EC 5.3.1.13). Next, A5P is coupled with phosphoenolpyruvate (PEP) and water to produce 3-deoxy-D-*manno*-octulosonate 8-phosphate (KDO8P) in a reaction catalysed by 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS) (EC 2.5.1.55). This is the first committed step toward formation of KDO. KDO8P, the product of the KDO8PS-catalysed reaction, is then dephosphorylated by KDO8P phosphatase (EC 3.1.3.45) to form KDO. KDO is then coupled with cytidine triphosphate (CTP) to produce the sugar-nucleotide cytidine monophosphate (CMP)-KDO catalysed by the enzyme CMP-KDO synthetase, prior to being transferred by KDO transferase to the acceptor lipid IV_A. At this point the secondary acyl chains of laurate and myristate are added in a stepwise fashion, producing the completed KDO_2-Lipid A (Figure 1.2).

This critical biosynthetic pathway for Gram-negative bacteria is a potential target for new antibacterial drug design and therapeutic strategies. In order to better understand these possibilities a firm understanding of the enzymes responsible is important. This thesis explores the workings of KDO8PS, the enzyme responsible for catalysing the first committed step

[i] In plants, KDO is found in the cell wall polysaccharides of green algae, and in the rhamnogalacturonan II pectin fraction of the primary cell walls of higher plants.[12,13]

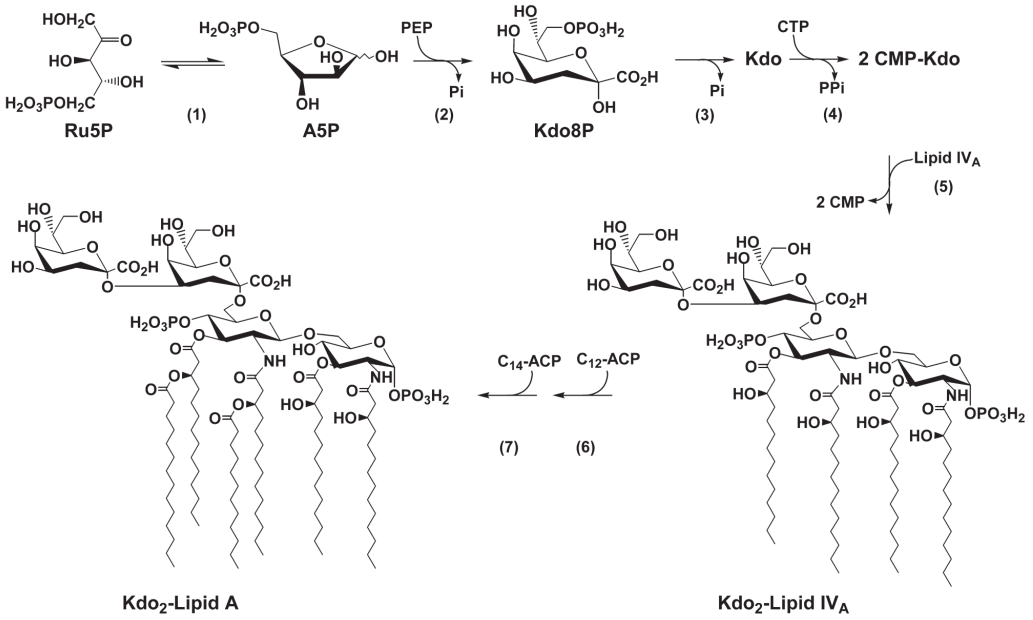


Figure 1.2: The biosynthetic pathway to the formation of KDO and the assembly of KDO_2-Lipid A.[7] The enzymes involved in the biosynthesis of KDO are (1) A5P isomerase, (2) KDO8PS, (3) KDO8P phosphatase, and (4) CMP-KDO synthetase. In *E. coli* two molecules of activated KDO are then sequentially transferred to lipid IV_A by (5) KDO transferase before the stepwise addition of the secondary acyl chains (6) laurate and (7) myristate. The figure is from Meredith et al. (2006).[7]

in the pathway. The remainder of this introductory chapter focuses on the structure, mechanism and function of KDO8PS, and how these attributes relate to a closely related enzyme.

1.2 The reaction catalysed by KDO8PS

The reaction catalysed by KDO8PS is the irreversible aldol-like condensation of the three-carbon compound PEP with the five-carbon phosphorylated aldose A5P (Figure 1.3).[15] During the reaction, C3 of PEP attacks C1 of A5P, producing an intermediate species that decays with the elimination of the phosphate derived from PEP, yielding inorganic phosphate and the eight-carbon sugar KDO8P as products.

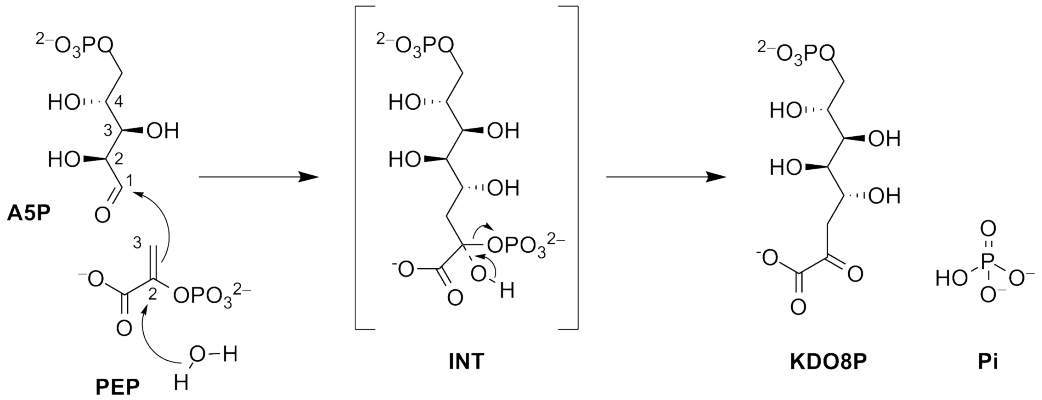


Figure 1.3: The aldol-like condensation reaction between PEP and A5P that is catalysed by KDO8PS.

It has been demonstrated in multiple studies[16–18] that during the reaction catalysed by KDO8PS, the *si* face of PEP reacts with the *re* face of the A5P aldehyde. The experiments in these studies used deuterated or fluorine-substituted PEP, and ^{1}H and ^{19}F nuclear magnetic resonance (NMR) spectroscopy to detect the stereospecifically-labelled KDO8P products. This facial selectivity has also been found to be favourable by computational studies.[19] Kinetics experiments have determined that KDO8PS operates an ordered sequential mechanism whereby the substrates, first PEP and then A5P, bind to the enzyme sequentially.[20] Product release consists first of inorganic phosphate and subsequently KDO8P. It has been determined that *Escherichia coli* KDO8PS (*Eco*KDO8PS) acts upon the acyclic form of A5P (and catalyses the reaction with 4-deoxyA5P in place of A5P, which is incapable of ring closure to the furanose form, with a similar k_{cat} value to that with A5P).[20] The rate-limiting step in the reaction (for *Aquifex aeolicus* KDO8PS (*Aae*KDO8PS)) is the formation of the intermediate ($k = 95\,\mathrm{s}^{-1}$), with breakdown of the intermediate to KDO8P having a five-times faster rate ($k = 500\,\mathrm{s}^{-1}$).[19] This is consistent with the rate of product formation.

The labile phosphate hemiketal intermediate (labelled INT in Figure 1.3) has been observed by mass spectrometry,[21–23] and modelled crystallographically.[24] Other mass spectrometry studies[25,26] have also ruled out the formation of a cyclic intermediate, confirming the formation of the acyclic intermediate.

The course of the reaction that leads to formation of this intermediate, and whether formation of the new C-C and C-O bonds is stepwise or synchronised, remains unclear.[27,28]

In the most chemically plausible scenario, C3 of PEP is thought to attack C1 of A5P, followed by attack of water giving rise to a transient oxocarbenium ion (Figure 1.4, Path A). However, if attack of water on C2 of PEP was to precede the reaction of PEP with A5P, then the reaction would instead proceed via a transient carbanion (Path B). The latter carbanion mechanism requires, however, unprecedented electrophilic character at C2 of PEP for reaction with a nucleophilic water molecule. Additionally, it is unlikely a resonance unstabilised carbanion would exist, with a p$K_a > 30$, inside an enzyme active site surrounded by chemical species with greater acidity. It is likely that such a carbanion would be very quickly quenched by hydrogen abstraction from a nearby species.[28] It seems most likely that the reaction between water, PEP and A5P is partially concerted, with C-C

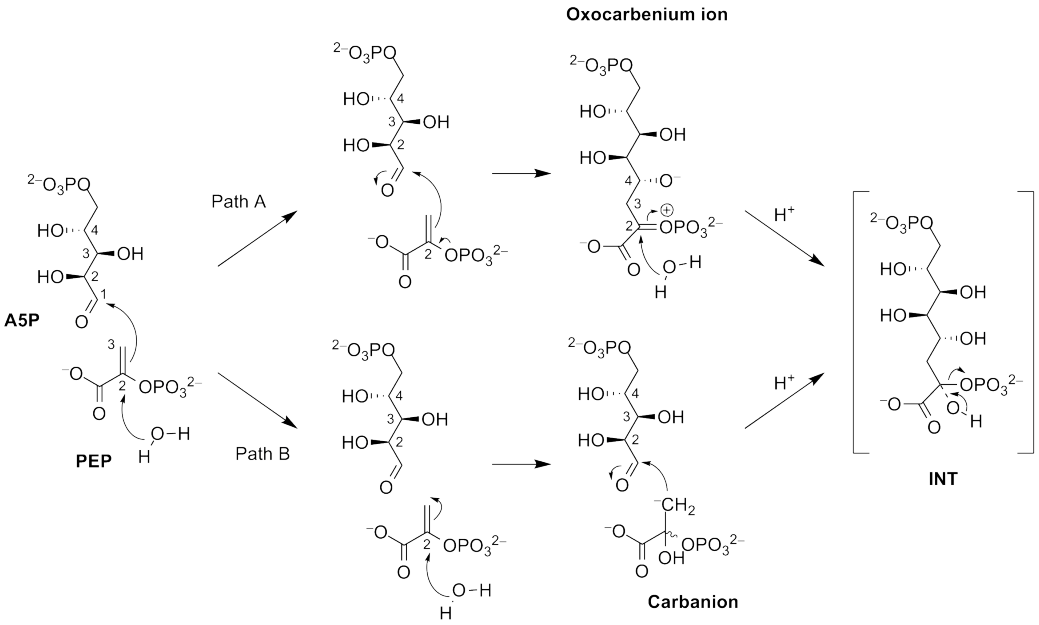
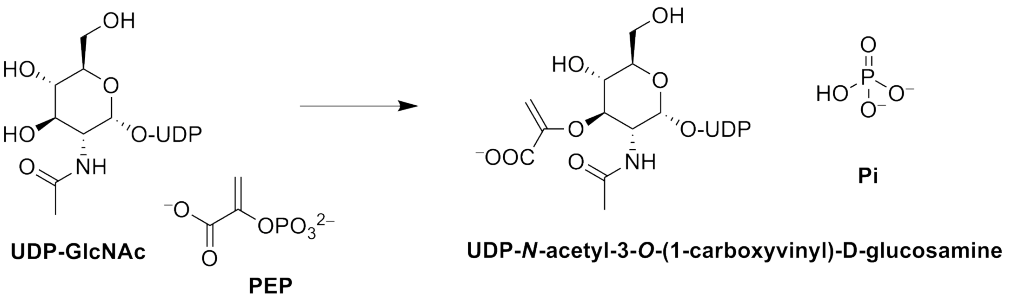


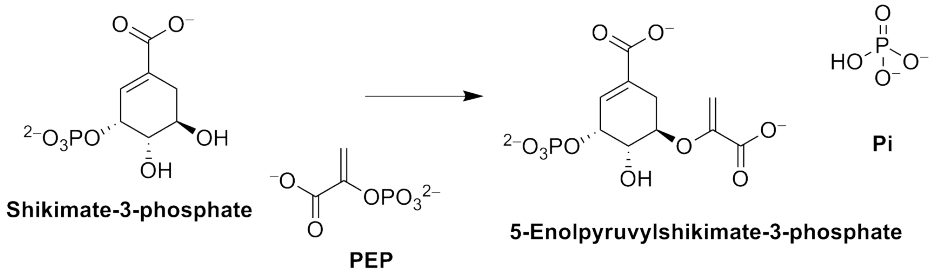
Figure 1.4: The two possible reaction pathways for the KDO8PS-catalysed reaction. Path A is through the formation of a transient zwitterionic oxocarbenium ion, and path B through a transient carbanion.

bond formation preceding C-O bond formation, a conclusion supported by computational studies.[19,29]

Independent of the true reaction pathway, both possible pathways will converge to a more stable tetrahedral intermediate, which will itself breakdown into KDO8P and inorganic phosphate. Unusually, the loss of phosphate is through the cleavage of the C-O bond of PEP, rather than the high-energy phosphate ester (P-O) bond common to other PEP-utilising enzymes.[22,30–32] This uncommon cleavage occurs in other enzymes too, such as UDP-*N*-acetylglucosamine enolpyruvyl transferase (Figure 1.5a)[33] and 5-enolpyruvoylshikimate 3-phosphate synthase (Figure 1.5b).[34] As observed for KDO8PS, these enzymes also catalyse the transfer of the enol ether of PEP to their respective alcoholic or aldehydic co-substrates, and share the same facial selectivity and reaction of an electrophile at C3 of PEP.[35] Furthermore, for each an oxocarbenium ion transition state or intermediate has been suggested



(a)



(b)

Figure 1.5: The reactions catalysed by enzymes that share a similar C-O bond cleavage to KDO8PS. (a) UDP-*N*-acetylglucosamine enolpyruvyl transferase. (b) 5-enolpyruvoylshikimate 3-phosphate synthase.

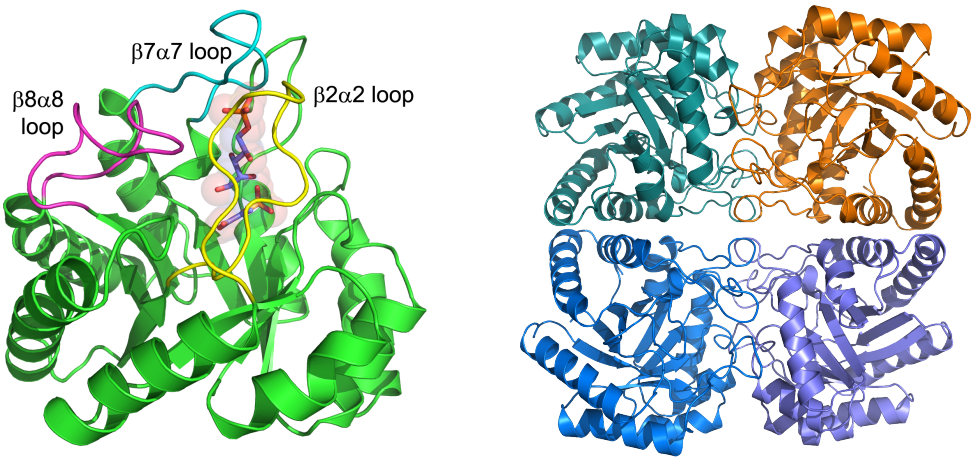
to be formed along the reaction coordinate. These precedents, together with kinetic and structural data of a mechanism-based inhibitor,[36] studies with PEP analogues[37] and intramolecular models[38] indicate an oxocarbenium-based reaction mechanism (whereby C-C bond formation precedes C-O bond formation) is likely to occur in KDO8PS.[28]

1.3 Structure of KDO8PS

The monomeric structure of KDO8PS is a common $(\beta/\alpha)_8$ triose phosphate isomerase (TIM)-barrel fold, enhanced only by extensions to the βα loops that extend from the C-terminal end of the barrel (Figure 1.6a). As with other TIM-barrel enzymes, the C-terminal region of the barrel forms the active site. KDO8PSs have been isolated and studied from a variety of sources: *E. coli*,[15] *Neisseria meningitidis*,[39] *A. aeolicus*,[40] *Salmonella typhimurium*,[11] *Aquifex pyrophilus*,[28] *Helicobacter pylori*,[41] *Arabidopsis thaliana*[42] and *Neisseria gonorrhoeae*.[43,44] In solution and crystal form the ≈30 kDa monomers form homotetramers (Figure 1.6b).[39,40,45–47] This tetrameric quaternary structure is the functional form of the enzyme, with each monomeric subunit containing an independent active site.

PEP binding is achieved by interactions with residues close to the ends of the β-strands of the core barrel, whereas the binding site for A5P protrudes further from the core of the barrel. Three long loops (β2α2, β7α7, β8α8), which link the ends of the eponymous β-strands and corresponding α-helices, extend from the barrel, create the A5P binding site and support inter-subunit contacts in the tetrameric protein. These loops and their specific active-site roles will be investigated and discussed in more detail later in this thesis.

There are two forms of KDO8PS, categorised according to the dependence on a metal ion for catalytic activity. Although the binding site for the metal ion is deep within the active site of the enzyme, close to the positions of the substrates PEP and A5P, both metal-dependent and metal-independent KDO8PSs are extraordinarily similar. Metal dependency and the two forms



(a) **(b)**

Figure 1.6: The structure of *Aae*KDO8PS (PDB code 2NX3). (a) A single monomer of KDO8PS showing the loops that extend the core barrel and the location of the two substrates within the active site. The β2α2 loop is coloured yellow, β7α7 loop cyan and β8α8 loop magenta. PEP and A5P are shown as both sticks and spheres with carbon atoms coloured purple. (b) The homotetramer of KDO8PS. Each monomer is uniquely coloured.

of KDO8PS are more extensively discussed in section 1.5 of this chapter.

1.4 The closely related enzyme DAH7PS

The enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH-7PS) (EC 2.5.1.54) catalyses the reaction between the four-carbon sugar D-erythrose 4-phosphate (E4P) and PEP, in a similar aldol-like condensation to KDO8PS (Figure 1.7). Compared to A5P, E4P is a four-carbon rather than five-carbon sugar and has the opposite configuration at C2. DAH7PS is responsible for catalysing the first committed step of the shikimate pathway, which gives rise to among many end products, the aromatic amino acids; tyrosine, phenylalanine and tryptophan.[48] The shikimate pathway, and hence DAH7PS, is present in only plants and bacteria, which makes DAH7PS a potential target for anti-bacterial drug design.[49] Unlike KDO8PSs, DAH7PSs

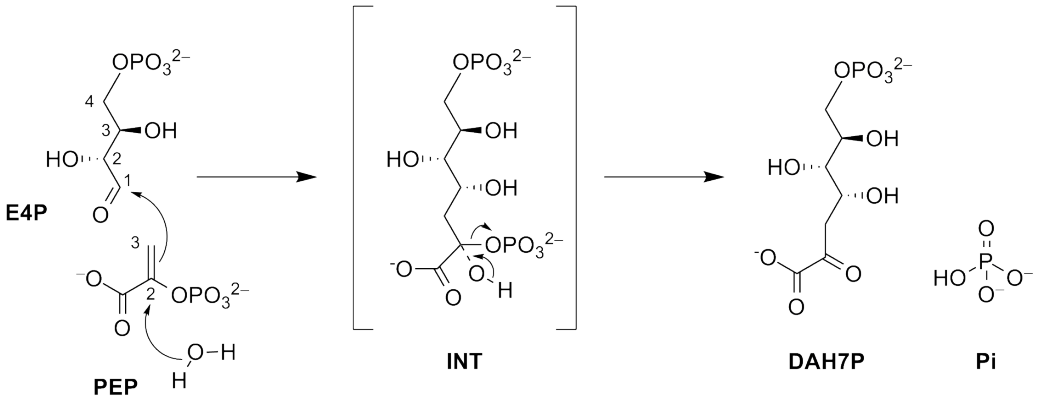
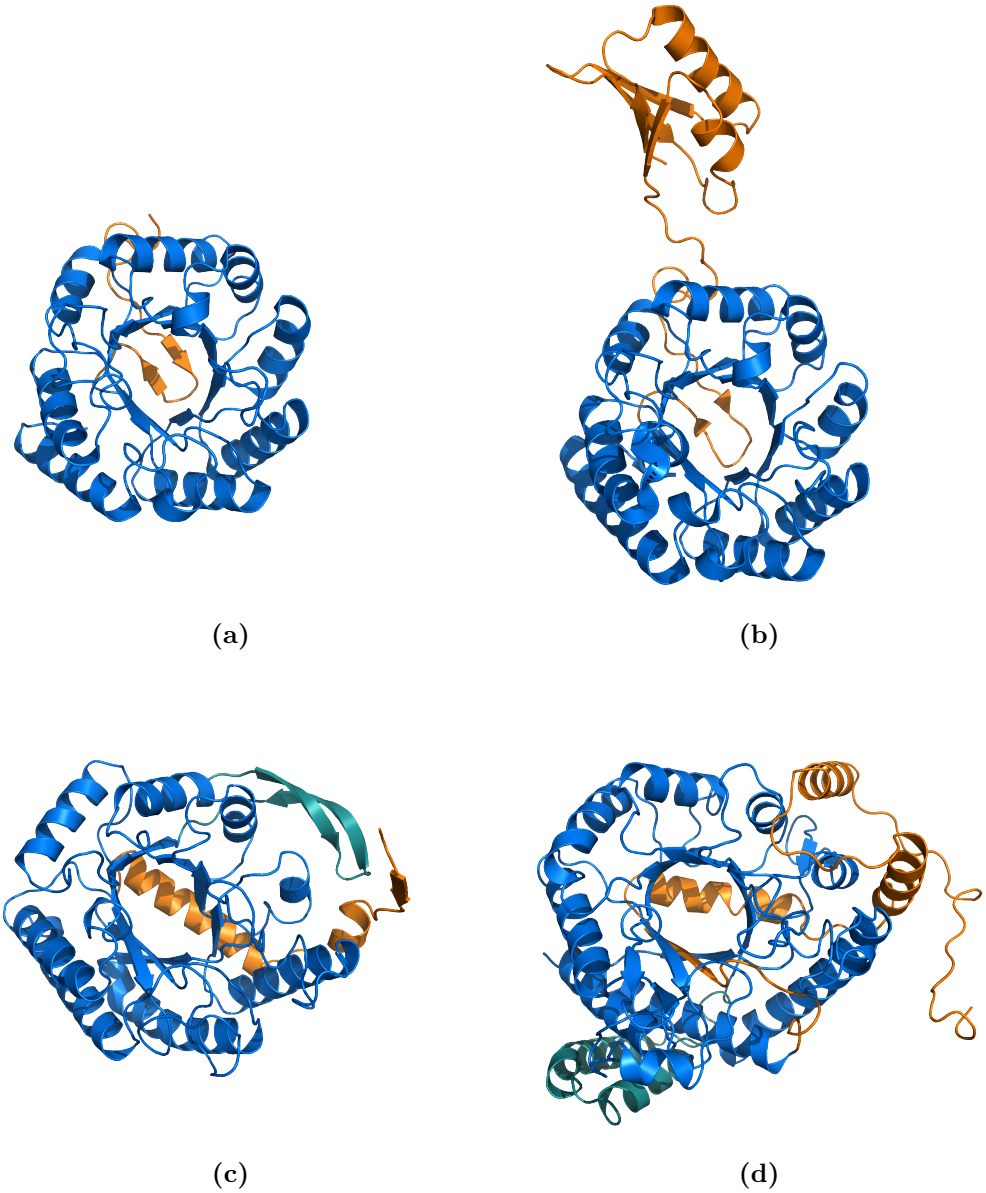


Figure 1.7: The aldol-like condensation reaction between PEP and E4P that is catalysed by DAH7PS.

are more structurally diverse, however all are likewise based on a core $(\beta/\alpha)_8$ TIM-barrel monomer fold (Figure 1.8). The structural diversity arises in the form of decorations to the core barrel: additions to the N- or C-terminal and extensions to the loops between β-strands and α-helices. Each DAH7PS can be classed as belonging to one of three different groups (Iα, Iβ or II) based on the nature of these additions, while the presence of the decorations correlate to, and are responsible for, enabling feedback inhibition.[50–54]

The type Iα DAH7PSs have both an N-terminal extension and an insertion between α5 and β6 composed of a two-stranded anti-parallel β-sheet (Figure 1.8c), which pairs with that of an adjacent monomer at dimer interfaces to create binding sites for tyrosine and phenylalanine.[53,55] Type II DAH7PSs always have an N-terminal extension (composed of a β-strand and two α-helices) and often an α2β3 loop extension (pair of α-helices), though this is sometimes absent (Figure 1.8d).[56] In this type, the additions to the core barrel also create binding sites for the aromatic amino acids.[54] The Iβ DAH-7PSs are the simplest proteins and have either an N- or C-terminal domain, or a small N-terminal extension to the core barrel structure (Figure 1.8a and 1.8b).[57,58] The N-terminal domain has been found to be an ACT/ferredoxin-like or chorismate mutase domain, while the C-terminal domain is found as a chorismate mutase domain. The presence of the ACT domain in *Tma*DAH7PS has been demonstrated to facilitate feedback inhibition by rearranging to



(a) **(b)**

(c) **(d)**

Figure 1.8: The monomeric subunit structure of various DAH7PSs. The common TIM-barrel fold core of each structure is coloured blue, N-terminal extensions and domains are coloured orange, and other barrel extensions are coloured green. (a) *Pfu*DAH7PS (type Iβ, PDB code 1ZCO), (b) *Tma*DAH7PS (type Iβ, PDB code 1RZM), (c) *Nme*DAH7PS (type Iα, not deposited), (d) *Mtu*DAH7PS (type II, PDB code 3NV8).

hinder active-site access, and is locked in place by tyrosine molecules.[58] The absence of any major barrel decorations is associated with a lack of feedback inhibition.[59] KDO8PSs are most similar to members of Iβ DAH7PS, being both undecorated and unregulated.

Inspection of the active sites of both DAH7PS and KDO8PS reveals that despite low sequence identity (≈30 percent between KDO8PS and Iβ DAH7PS), the active-site architectures of both enzymes are very similar, composed of many shared (and hence conserved) residues. The metal-binding site (*vide infra*) is conserved, and PEP is bound in a similar position. E4P has been modelled to sit in the active site similarly to A5P in KDO8PS, with the aldehyde closest to PEP, and the phosphate moiety most distal.[59] Like KDO8PS, the reaction catalysed by DAH7PS has the same facial selectivity (*re* face of E4P and *si* face of PEP react together),[60,61] DAH7PS operates the same ordered-sequential mechanism,[62] and the equivalent acyclic intermediate is formed and decays with the loss of phosphate through cleavage of the C-O bond.[63] Unlike for KDO8PS for which both metal-dependent and metal-independent forms have been described, all characterised DAH7PSs are metal-dependent.

The similarity between KDO8PS and DAH7PS is clearly demonstrated structurally (Figure 1.9) and by the similarity in substrates and catalysed reaction chemistry. Unsurprisingly, KDO8PSs (and the divergent types of DAH7PS) have been proposed to have evolved from a common Iβ-like ancestor (Figure 1.10).[45,64] This ancestral enzyme is expected to have had promiscuous activity toward a range of phosphorylated aldose substrates, been metal-dependent, and been insensitive to feedback inhibition. *Pfu*DAH7PS has all these characteristics and has been suggested to resemble this ancestor.[59]

1.5 Metal dependency of KDO8PS

KDO8PSs can be divided into two distinct classes by whether they are either dependent on or independent of a divalent metal ion for catalytic activity. In

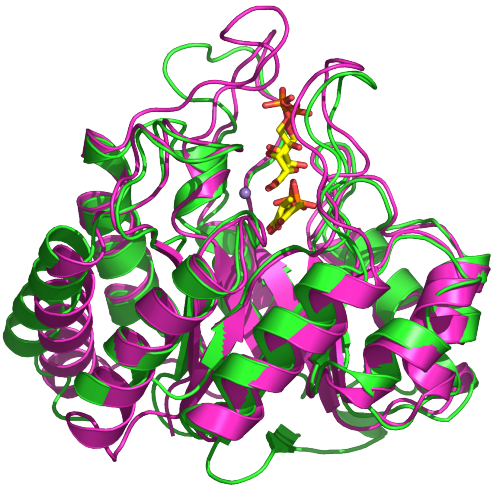


Figure 1.9: A structurally aligned superposition of the *Pfu*DAH7PS monomer (coloured green, PDB code 1ZCO) and the engineered metal-independent *Aae*-KDO8PS monomer (coloured pink, PDB code 2NX3). The carbon atoms of PEP, E4P and A5P are coloured yellow in both structures, and the Mn^{2+} in *Pfu*DAH7PS is coloured purple.

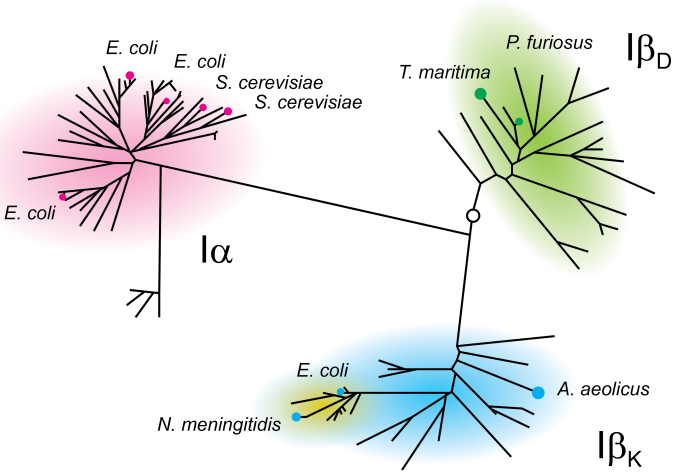


Figure 1.10: Phylogenetic tree showing the homology between the type Iα and Iβ DAH7PSs, including the Iβ-like KDO8PSs. The metal-independent KDO8PSs are highlighted yellow. *Afe*KDO8PS is not included in the tree. The open circle represents the position of the suggested Iβ-like ancestral protein. Redrawn from Jensen et al. (2002).[64]

metal-dependent KDO8PSs, an active-site divalent metal ion is positioned close to the carboxylate of PEP and the aldehyde of A5P, held in place by four conserved residues: an Asp, Glu, Cys and His (Figure 1.11). *In vitro*, metal-dependent KDO8PSs are typically maximally active in the presence of the divalent ions of Cd, Mn and Co, with varying lower levels of activity in the presence of Ni, Fe, Cu, Ca and Zn, and lower yet activity with other divalent metal ions.[40,41,47,65,66]

In metal-independent KDO8PSs the residues of the metal-binding site are retained and equally conserved, except for one of the four metal-ligand residues (Figure 1.11). The fourth metal ligand, Cys, is substituted for by a conserved Asn residue. The metal dependency of uncharacterised KDO8PSs can be putatively assigned by inspection of the amino acid sequence for the presence of the Cys or Asn residue, which in crystal structures of KDO8PSs is on the β1α1 loop.[45,64] Other than this single substitution, there are no significant differences between the active sites of both types of KDO8PS.

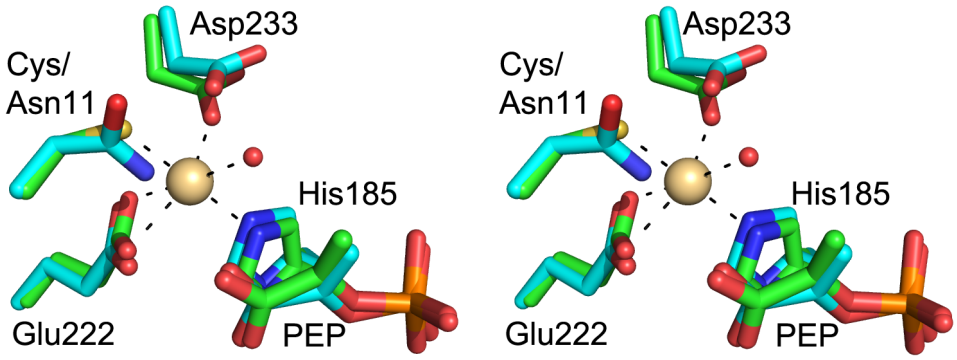


Figure 1.11: Stereoview showing a superposition of the metal-binding sites of wild-type and engineered metal-independent *Aae*KDO8PS. The carbon atoms of the wild-type *Aae*KDO8PS structure (PDB code 1FWW) are coloured green and Cd^{2+} is coloured wheat. The carbon atoms of the engineered metal-independent *Aae*KDO8PS structure (PDB code 2NX3) are coloured cyan.

1.5.1 Interconversion of metal dependency

It is possible to interconvert the metal dependency of different forms of KDO8PS. Metal-dependent KDO8PSs have been converted to gain metal-independent activity, and vice versa (Figure 1.12).[24,28,39,67,68] Although in most cases Asn-to-Cys substitution is sufficient to achieve metal activation, or, metal-independent activity (Cys-to-Asn), in no case has the single substitution been sufficient for the mutated enzyme to retain 100 percent of its native activity. Nor, when converting from a metal-independent to metal-dependent form, is the single substitution enough to eliminate metal-independent activity.

In the conversion of the metal-independent *Eco*KDO8PS, the requisite N26C mutation was coupled with M25P,[68] a mutation that represents an almost conserved difference between the two classes of KDO8PS. This was necessary to better restore the activity (k_{cat}) of the mutated enzyme, in the presence of Mn^{2+}, to that of the metal-independent wild-type enzyme. Obligate metal dependency was created in the interconversion of the likewise

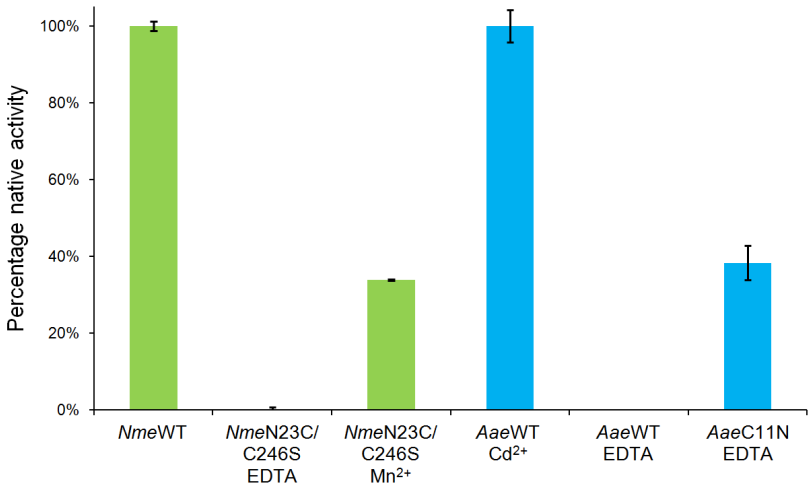


Figure 1.12: Relative activity (computed from k_{cat} values) in the presence of the most-activating divalent metal ion or EDTA of wild-type and metal-dependency-converted mutants of *Nme*KDO8PS and *Aae*KDO8PS. *Nme*KDO8PS is natively metal-independent and was converted to become metal-dependent,[39] and *Aae*-KDO8PS (metal-dependent) was made metal-independent.[24,67]

metal-independent *Nme*KDO8PS, only by also mutating another residue (C246S) together with N23C.[39] Left unchanged, Cys246 in the N23C mutant of *Nme*KDO8PS forms an undesirable disulfide bond with the introduced Cys residue. For the natively metal-dependent *Aae*KDO8PS, a metal-independent version was created with the single Cys-to-Asn mutation (C11N) and this mutation was coupled with two other mutations (S235P and Q237A).[24,67] The respective Pro and Ala residues were identified as being conserved in metal-independent KDO8PSs, and the identity of these positions as Ser and Gln in wild-type *Aae*KDO8PS was noted as being important for optimising metal coordination.

These observations suggest metal-independent KDO8PSs have evolved compensatory changes to better facilitate metal-independent activity, and therefore engineered metal-dependent KDO8PSs probably represent evolutionary intermediates.[68] It is believed the metal-independent KDO8PSs are the evolutionarily youngest in the larger DAH7PS family, and that the metal-dependent KDO8PSs evolved first, from an ancestral Iβ DAH7PS-like enzyme.[64] Notably, substitution of the metal-ligand Cys-to-Asn in *Pfu*DAH7PS is not capable of removing the metal-ion dependency.[69] Indeed, the metal ion appears to be required for catalytic activity in all DAH7PSs,[56,57,70,71] consistent with all known sequences having a Cys at the appropriate position.

1.5.2 Role of the metal ion

It was originally proposed that the KDO8PS-catalysed reaction may proceed through a different order of steps depending on the metal dependency of the enzyme.[28,36,65] For metal-dependent KDO8PSs it was proposed the reaction could proceed through the carbanion mechanism (Figure 1.4), whereby the metal ion activates a water molecule for attack on C2 of PEP, or during the reaction the metal ion helps to stabilise the incipient negative charge on the aldehyde of A5P (the two roles being mutually exclusive). The latter, with the carbonyl coordinated to the metal ion, is likely the role the indispensable metal ion plays in DAH7PS.[69] However, the metal ion in KDO8PSs can

be dispensed by a single mutation, although at the expense of a reduction in activity relative to wild-type enzyme. Additionally, metal-independent activity is not extinguished in engineered metal-dependent enzymes; these enzymes tend to exhibit only increased activity in the presence of a metal ion, rather than an obligate dependence. Therefore the role of the metal ion seems unlikely to be to activate the aldehyde or a water molecule for attack toward PEP, indicating mechanistic commonality between metal-dependent and metal-independent KDO8PSs.[28] The mechanistic similarity between both forms of KDO8PS was also demonstrated by experiments using (*E*)- and (*Z*)-fluorinated PEP (phosphoenol-3-fluoropyruvate), where both forms of KDO8PS had the same high selectivity for (*E*)- over (*Z*)-fluoro-PEP.[18] This selectivity was not observed for DAH7PS.

It would seem that rather than being directly involved in the chemistry of the catalysed reaction, the metal ion more likely has an important structural role to play, perhaps, by helping to secure the active site and properly orient the substrates and active-site residues.[28,29] In the absence of the metal ion, this role is partially fulfilled by the Asn residue that replaces the metal-ligand Cys and is complemented by other compensatory changes, particularly related to securing in-place this area of the active site. It is unclear why in KDO8PS the role of the metal ion has changed in comparison to DAH7PS.

1.5.3 Position of reactive active-site water molecule

Whereas the key features of the enzyme mechanism for KDO8PS are known, the chirality of the tetrahedral intermediate is unclear. The water molecule that reacts with PEP must attack from one side or the other with a defined geometry. However, it is presently unclear where the water that reacts with C2 of PEP is positioned and hence whether it attacks the *re* or *si* face of PEP. It has been proposed that the reactive water is on the *si* face (closest to the metal-binding site) of PEP, leading to overall syn-addition of water and A5P to PEP for the reaction.[24]

A water molecule is often observed on the *si* face of PEP in crystallo-

graphic models of *Aae*KDO8PS (Figure 1.13), but is less commonly observed when both PEP and A5P bound. It has been suggested that a water molecule may not be present at the beginning of the reaction but may enter the active site from the bulk phase after formation of the oxocarbenium intermediate (should it be sufficiently stable) and react with the *si* face of C2 of the intermediate.[24] The acyclic intermediate has also been modelled in the crystal structure of an engineered metal-independent *Aae*KDO8PS, where the tetrahedral configuration of the intermediate suggests reaction with water on the *si* face.[24] However, the electron density for the hydroxyl group derived from this water in the intermediate is unconvincing (Figure 1.15c). The only other evidence favouring the *si*-face water is computational studies.[19] These, unfortunately, calculated potential energy surfaces without complete consideration of the active-site environment in which catalysis takes place, and therefore its potential contribution.

It has also been suggested that a water molecule on the *si* face of PEP in metal-dependent KDO8PSs is a metal ligand, and as such the metal ion would activate the water molecule for reaction with PEP.[27,72] Interpretation of quantum mechanical calculations has suggested the metal-ligand Asp may be responsible for deprotonating the metal-ligand water to initiate catalysis.[73] This is at odds with the binding of A5P, and would also seem to favour the

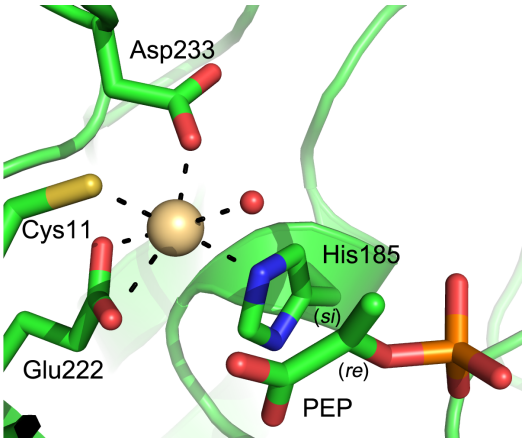


Figure 1.13: The metal-binding site of *Aae*KDO8PS (PDB code 1FWW) showing the metal-bound water molecule on the *si* face of PEP.

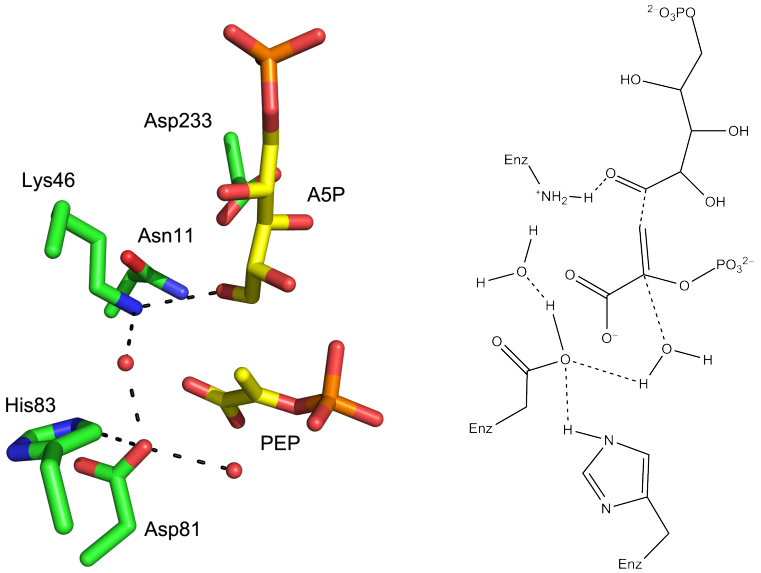
less-likely carbanion reaction pathway.

How a water molecule on the *si* face of PEP is activated in both metal-independent and engineered metal-independent KDO8PSs has not yet been directly addressed. Whatever the position of the water molecule, the mechanism in which its reaction with PEP is facilitated, must only occur when both PEP and A5P are bound to the enzyme, as in the absence of A5P there is no wasteful breakdown of PEP.

Anti-addition of water to (the *re* face of) PEP is also possible. This would presumably be the more favourable mechanism and is supported by the observation of a water molecule on the *re* face of PEP in many crystallographic models, including those from both metal-dependent and metal-independent KDO8PSs.[24,39,72] Unlike the water molecule on the *si* face, in most structures of *Aae*KDO8PS with both PEP and A5P bound, a water molecule is observed on the *re* face of PEP.

It has been suggested that the *re*-face water is stabilised by a hydrogen-bonding chain that includes a conserved Asp and His residue.[72] The chain could transfer a proton from the water to Asp and ultimately to His (a more likely final acceptor). The equivalent residue to the Asp is conserved as a Glu in DAH7PS, however there is no equivalent residue to the conserved His.

It has been proposed for DAH7PS, that the equivalent nucleophilic water molecule is on the *re* face of PEP, and deprotonation of it is linked via a proton relay chain to protonation of the aldehyde group of E4P.[74] This relay is through the conserved Glu mentioned above, another water molecule, and a conserved Lys residue. A water molecule on the *re* face of PEP has been observed in crystallographic models of different DAH7PS types from a variety of species.[74–77] The residues and water molecule of the DAH7PS proton relay are equally conserved in KDO8PS, except with the substitution of the Glu for Asp (Figure 1.14). Given the homology between the two enzymes, it is conceivable KDO8PS may operate a similar scheme. Therefore, because of the ease with which the metal ion can be dispensed and the strong evolutionary ties KDO8PS has to DAH7PS, it is tempting to suggest that the reactive



(a) (b)

Figure 1.14: The proton relay chain that may link deprotonation of the nucleophilic water molecule on the *re* face of PEP with protonation of the carbonyl oxygen of A5P. (a) Crystal structure model of *Aae*KDO8PS (PDB code 2NX3) showing the interactions from the *re*-face water molecule to the A5P carbonyl. The carbon atoms of PEP and A5P are coloured yellow. (b) Possible interactions involved in catalysis.

water molecule is most likely on the *re* face of PEP.

1.6 A5P binding and substrate specificity

Crystallographic models with PEP and A5P bound to the enzyme have been solved for the thermophilic *Aae*KDO8PS. In wild-type and mutant models of *Aae*KDO8PS with both PEP and A5P bound, A5P is modelled with its aldehyde closest to the PEP phosphate (Figure 1.15a).[72,73,78,79] While in crystallographic models of engineered metal-independent *Aae*KDO8PS, although in some cases the electron density is poor, the aldehyde of A5P is consistently positioned closest to (what was) the metal-binding site (Figure 1.15b).[24] The two orientations are loosely related by a 180 degree rotation

about the long axis of A5P, but both still present the *re* face of A5P to PEP. The structure of *Aae*KDO8PS has also been solved in the presence of ribose 5-phosphate (R5P) (C2 epimer of A5P, which is not a substrate for *Aae*KDO8PS, *vide infra*).[27] In this model, the aldehyde is located toward the PEP phosphate.

Although several structures of *Aae*KDO8PS have been solved in which A5P is bound to the active site, the binding mode of A5P in the active site of KDO8PS, specifically that which is catalytically competent, is presently unclear. While it is accepted that the phosphate moiety of A5P binds in the active site most distal to PEP, and the aldehyde of A5P is bound closest to PEP, the direction in which the aldehyde points is ambiguous. There is evidence to suggest the aldehyde points toward the position of the metal-binding site (or remnants of it) and PEP carboxylate (Figure 1.15b), and that it is toward the position of the PEP phosphate (Figure 1.15a).

The tetrahedral reaction intermediate has also been modelled in crystal structures of engineered metal-independent *Aae*KDO8PS and the product (KDO8P) in wild-type *Aae*KDO8PS (Figure 1.15c and 1.15d).[24,73] In both cases, the hydroxyl derived from the aldehydic oxygen of A5P is pointing in the direction of the metal ion and PEP carboxylate. An A5P and PEP mimicking inhibitor[ii] has also been crystallographically modelled bound to *Aae*KDO8PS, also with what is equivalent to the A5P aldehyde closest to the metal-ion site.[27]

1.6.1 Mechanistic implications of A5P binding mode

The position of A5P in the active site must be considered in regard to mechanistic implications. In the catalysed reaction, the nascent negative charge on the aldehyde of A5P must be quenched by abstraction of a proximal hydrogen ion. Additionally, for reaction at C1 of A5P, its position in the active site must be precisely controlled. Careful positioning of this reactive

[ii] {[(2,2-dihydroxyethyl)(2,3,4,5-tetrahydroxy-6-phosphonooxyhexyl)amino]methyl}-phosphonic acid

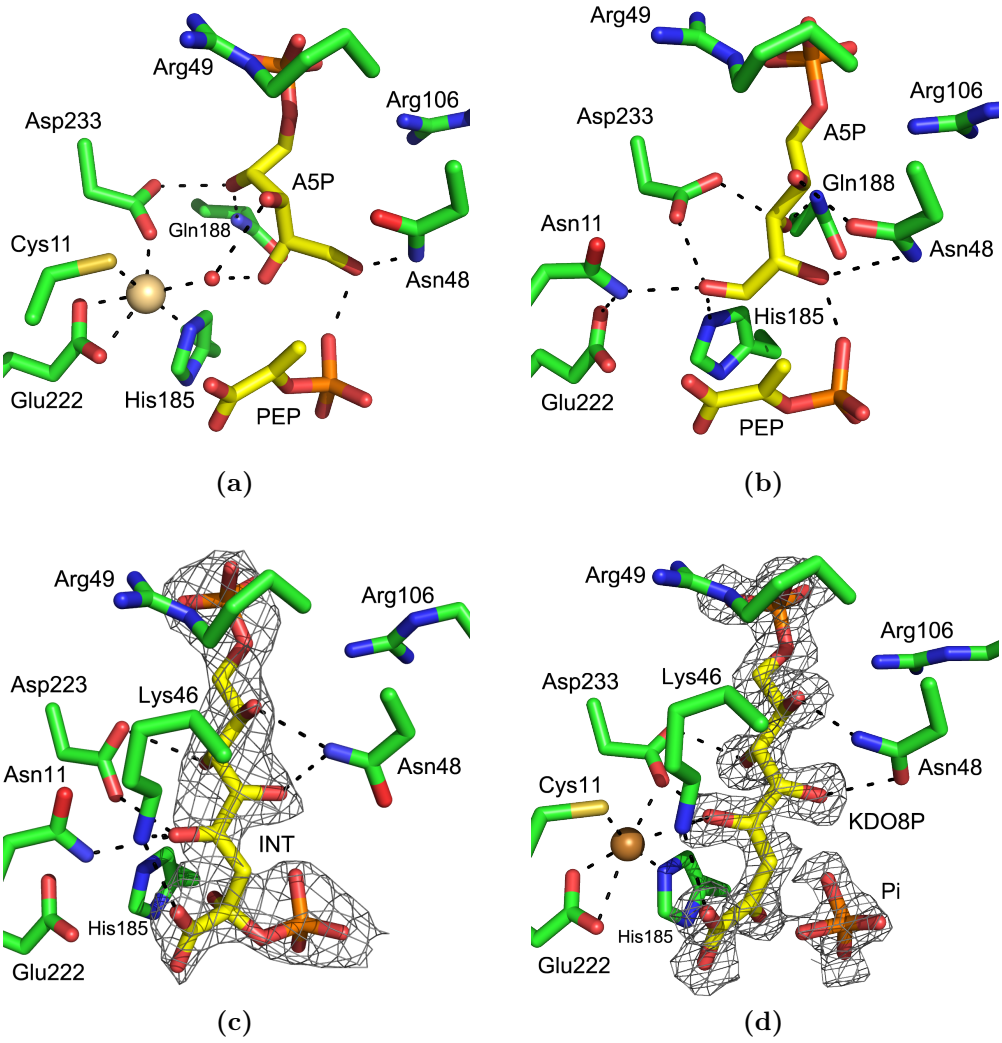


Figure 1.15: The binding modes of A5P, tetrahedral reaction intermediate and products to *Aae*KDO8PS. (a) PEP and A5P bound to wild-type *Aae*KDO8PS (PDB code 1FWW). The carbonyl of A5P is pointing toward the phosphate moiety of PEP. The metal ion is Cd^{2+}. (b) PEP and A5P bound to the engineered metal-independent *Aae*KDO8PS with the carbonyl of A5P toward what was the metal-binding site (PDB code 2NX3). (c) The tetrahedral reaction intermediate modelled in the engineered metal-independent *Aae*KDO8PS (PDB code 2NX3). (d) The products KDO8P and inorganic phosphate modelled bound to Cu^{2+}-*Aae*-KDO8PS (PDB code 3E12). Protein carbons are coloured green and those of ligands yellow. Electron density maps are contoured at 1 σ.

end of the A5P molecule would at the very least involve both the aldehyde and C2-OH having well-conserved interactions with nearby active-site residues or the co-substrate PEP (Figure 1.16a).

If the A5P aldehyde were to point towards the position of the PEP phosphate, then it has been proposed that a monoanionic (rather than dianionic) phosphate moiety of PEP could supply the proton required to capture the developing negative charge.[69] Interestingly, in KDO8PS, there is a conserved Phe in the PEP-phosphate binding pocket, replacing a conserved Arg in DAH7PS, which may favour the PEP phosphate to be in the monoanionic form (Figure 1.16b). This binding mode is consistent with the ability to interconvert the metal dependency of KDO8PSs, and has C2-OH of A5P toward the metal-binding site, securing the position of C1. Alternatively, if the A5P aldehyde were to point toward the position of the metal-binding site and PEP carboxylate, then it would be tempting to suggest, that like in DAH7PS, the metal ion directly activates the aldehyde to facilitate attack by PEP. However, given the ease with which metal dependency can be interconverted this direct

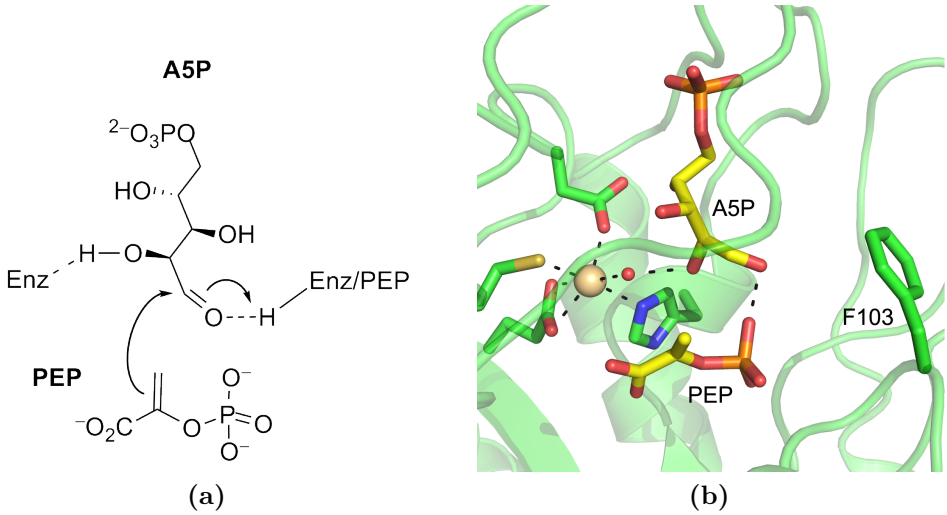


Figure 1.16: (a) Important interactions with A5P for precise positioning in the active site for reaction and supply of a proton to the nascent negative charge on the aldehyde group. (b) The structure of *Aae*KDO8PS (PDB code 1FWW) showing the interactions of the aldehyde group and C2-OH of A5P. The carbon atoms of PEP and A5P are coloured yellow, and Cd^{2+} is coloured wheat.

involvement seems unlikely to occur in KDO8PS. If A5P adopts this position, activation of the aldehydic oxygen must be achieved by other means.

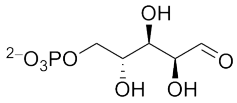
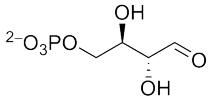
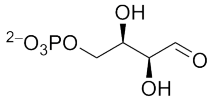
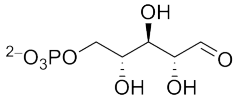
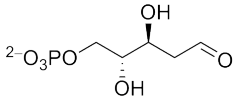
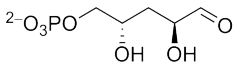
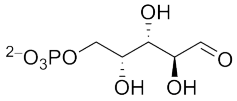
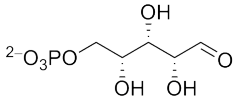
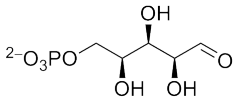
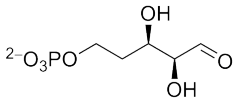
It is interesting to note that regardless of the mode of A5P binding that is catalytically competent, multiple binding modes have been observed. Molecular dynamics simulations of A5P binding in a C11N mutant of *Aae*-KDO8PS have likewise found no preferential mode of A5P binding.[19] Accurate interpretation of how real or useful the various observed modes are, and how representative they are of all KDO8PSs is limited by the dual substrate, intermediate and product bound structures being all of only *Aae*KDO8PS.

1.6.2 Substrate specificity

Unsurprisingly, given the apparent need to carefully arrange A5P in the active site, without an absolute requirement for any direct involvement of a metal ion, KDO8PS has specific substrate configuration requirements; the carbon atoms that bear the hydroxyl groups that decorate A5P must have particular configurations for both catalysis to occur at all, and to enable efficient catalysis. Substrate analogues of A5P, based on both the removal of hydroxyl groups and changes to their configuration have been tested on KDO8PSs from different species (Table 1.1). These results have clearly indicated that the presence of hydroxyl groups at C2 and C3 of A5P are critical for catalysis, while the hydroxyl group that decorates C4 of A5P is less important.[15,20,27,40,44,69,80–82] This contrasts with the more promiscuous substrate specificity exhibited by DAH7PS, whereby the configuration of the hydroxyl groups is largely unimportant, and catalysis is also possible with five-carbon phosphorylated aldoses such as A5P.[59,69,80,83]

A5P can be viewed as being held in place by two areas of key interactions at opposite ends of the molecule. Firstly, as discussed, the interactions of the aldehyde group, C2- and C3-OH can be presumed to precisely orient the reactive end of A5P close to PEP. Secondly, the phosphate moiety at the other end of A5P interacts with many conserved residues that line this region of the active site. But the specific interactions the hydroxyl groups of A5P

Table 1.1: Alternative substrates for KDO8PS.

Phosphorylated aldose		Substrate
	A5P	Yes
	E4P	No[a,b,c,d]
	T4P	No[a]
	R5P	No[a,b,c]
	2-deoxyR5P	Poor[a,b]
	3-deoxyA5P	No[b]
	L5P	No[a]
	D-X5P	No[a]
	L-X5P	Yes[a]
	4-deoxyA5P	Yes[b]

A5P: D-arabinose 5-phosphate, E4P: D-erythrose 4-phosphate; T4P: D-threose 4-phosphate; R5P: D-ribose 5-phosphate, L5P: D-lyxose 5-phosphate; X5P: xylose 5-phosphate. [a]*Nme*KDO8PS, [b]*Eco*KDO8PS, [c]*Aae*KDO8PS, [d]*Ngo*KDO8PS.

make with the residues that line the active site, and that of the A5P carbonyl itself, depend on the A5P orientation.

In substrate-bound crystal structures of *Aae*KDO8PS, both of wild type and engineered metal independent, the A5P-enzyme interactions are varied. When the A5P carbonyl is toward the PEP phosphate (wild-type *Aae*-KDO8PS, Figure 1.15a) it has interactions with the PEP-phosphate moiety, a water molecule and a conserved Asn residue. C2-OH of A5P interacts also with the PEP phosphate and a metal-ligand water molecule. C3-OH interacts with the same conserved Asn residue as the carbonyl, and the same metal-ligand water molecule as C2-OH. C4-OH interacts with the metal-ligand Asp, and a conserved Gln near the beginning of the β7α7 loop. With the A5P carbonyl toward the metal-binding site, observed in the engineered metal-independent *Aae*KDO8PS, and in modelled intermediates and product in the same enzyme (Figures 1.15b–1.15d), then the interactions are different. Here, the carbonyl interacts with the Cys-substituting Asn, a conserved Lys (which as discussed potentially delivers a proton), the PEP carboxylate, and the metal-ligand Asp. C2-OH interacts with the PEP phosphate, a conserved Asn residue and a water molecule. The C3-OH interacts with the metal-ligand Asp, and the conserved Gln on the β7α7 loop. C4-OH interacts, like C2-OH, with a conserved Asn residue. In both orientations, the carbonyl, C2-OH and C3-OH of A5P interact in different combinations with the same conserved active-site features, and the substrate specificity exhibited by KDO8PS can be equally rationalised. It is therefore unclear from the known substrate specificity which of the two observed A5P orientations is catalytically competent.

1.7 Goals of this thesis

The greater DAH7PS family of enzymes, which includes KDO8PS, represent an example of divergent evolution where an ancestral enzyme with broad substrate specificity has evolved over time into a range of similar but distinct enzymes. The objective of this thesis was to learn what defines a KDO8PS,

with a particular focus given to the mechanisms by which substrate specificity and selection is achieved. More generally, it was to also understand the roles of the subtle differences in active-site architecture between DAH7PS and KDO8PS and the effect of these on the way in which catalysis takes place. This knowledge was ultimately used to inform rational mutagenesis in an attempt to "un-evolve" a KDO8PS back into a DAH7PS.

The specific goals of the thesis were to:

- Use bioinformatic analyses to relate sequence and structure and uncover key differences between DAH7PS and KDO8PS;
- Investigate the metal dependency of a mesophilic metal-dependent enzyme that has metal-independent sequence features;
- Determine the role of a conserved substrate binding motif; the role it plays in substrate selection and specificity, and how it relates to the equivalent (but different) motif in DAH7PS;
- Probe the role of the conserved and extended β7α7 loop (that has substrate interactions) that is always present in KDO8PS and never present in DAH7PS;
- Look at the role of a conserved active-site Asp and its potential role in active-site proton transfer;
- Explore a conserved dimer interface area that differs between KDO8PS and DAH7PS, and its contribution to substrate selection;
- Combine the individually characterised differences in an attempt to modify the substrate specificity of KDO8PS to be more DAH7PS-like.

The experiments were performed using metal-dependent and metal-independent KDO8PS enzymes as a platform to identify potential differences between these two forms of KDO8PS. The metal-dependent enzyme used was *Afe*KDO8PS. This protein is from a mesophilic organism, and although being metal dependent, has some sequence features that tend only to occur in metal-independent KDO8PSs. This enzyme has previously been cloned

and initially characterised in our laboratory.[84] *Nme*KDO8PS was the metal-independent enzyme used in these studies. It is an enzyme that has previously been characterised in our laboratory,[39] again from a mesophilic organism, and also has known crystallisation conditions.

Chapter 2

Exploring enzyme evolution through analysis of sequence and structure

2.1 Introduction

The primary functional difference that distinguishes DAH7PS from KDO8PS is the choice of phosphorylated aldose substrate used in the reaction they each catalyse. DAH7PS catalyses the aldol-like condensation of PEP with E4P, whereas KDO8PS catalyses the same reaction, but utilising in much greater preference than vice versa, A5P in place of E4P. As has already been discussed in the introductory chapter, the two enzymes are evolutionarily related and very similar, in protein structure and enzyme mechanism. Given that function is a product of structure and that structure is a consequence of sequence, it follows that there must be identifiable sequence differences that account for the defining primary functional difference. Therefore it makes sense to analyse the amino acid sequence of each protein in conjunction with structural interpretation, as a method to highlight what sequence differences are responsible for the observed differences in function.

The amino acid sequence for a protein is variable between different species. Multiple sequence alignments (MSAs) provide a way to compare the variations and are useful to highlight the similarities between sequences. Sequence positions commonly occupied across sequences from different species by the same or similar amino acids have been selected for and maintained during evolution. These conserved positions or regions (motifs), must therefore be in some way important to an enzyme and its function. Establishing the differences in the patterns or profiles of conservation that fingerprint sequences of one protein, as compared to that of another, should then in theory yield the amino acid sequence differences responsible for the differences in function.

In this chapter the sequences of many KDO8PSs and Iβ DAH7PSs were collated and analysed by MSAs and hidden Markov model (HMM) profiles. The results of these sequence analyses are interpreted within the context of crystal structure models of *Aae*KDO8PS and *Pfu*DAH7PS.

2.2 Comparison between DAH7PS and KDO8PS

2.2.1 Multiple sequence alignments

Approximately 2069 annotated DAH7PS (of all types) and 1100 annotated KDO8PS sequences were downloaded from BRENDA.[85] A sequence-parsing script was written in Ruby and used to identify and separate 375 Iβ DAH7PS sequences (the type of DAH7PS most similar to KDO8PS) from the larger data set of DAH7PS sequences. MSAs of the Iβ DAH7PS sequences, and of the KDO8PS sequences were independently generated using ClustalX.

2.2.2 Profile hidden Markov models

While a MSA will identify positions and regions of sequence conservation, it is difficult to compare the conservation between two separate alignments. To facilitate this comparison and highlight the conservation differences between KDO8PS and DAH7PS, profile HMMs were computed for each MSA. Profile HMM analyses have previously been used to study co-evolving residues on the overall stability of KDO8PS.[86]

A profile HMM contains scores for the likelihood of each amino acid at each position in the MSA, which can be used to compute the relative entropy at each position. The relative entropy can be thought of as a measure of conservation or variability. Profile HMM were created for Iβ DAH7PS and (all) KDO8PS sequences (Figure 2.1). The differences in conservation between two compared profile HMM were established by calculating the magnitude of the difference in relative entropy at each position.

If the amino acid probabilities at each sequence position are known, then inherent within is a measure of how conserved that position is for a particular amino acid. Moreover, a core probabilistic model sequence can be derived from a profile HMM, where the amino acid at each position is that with the highest probability. The derived sequence, factored with conservation probabilities can be aligned and output in a FASTA-like format. The profile HMM of Iβ DAH7PSs was aligned with that of (all) KDO8PSs[i] and the aligned core probabilistic model sequence, annotated with MSA-derived conservation, was output (Figure 2.2). This profile-profile alignment of Iβ DAH7PS with KDO-8PS was used to re-index each sequence position number on a common basis. These position numbers are used throughout this chapter. An alignment of the most-likely KDO8PS sequence against those of some commonly referred

[i] While there are both metal-dependent and metal-independent forms of KDO8PS, and all Iβ DAH7PSs are metal-dependent enzymes, there are only minimal differences between the two forms of KDO8PS (*vide infra*). Hence, using sequences of both forms of KDO8PS in the comparison with Iβ DAH7PS is unlikely to confuse the identification of differences between the two enzymes, and the identification of the most important sequence features will be advantaged by analysing a greater number of sequences.

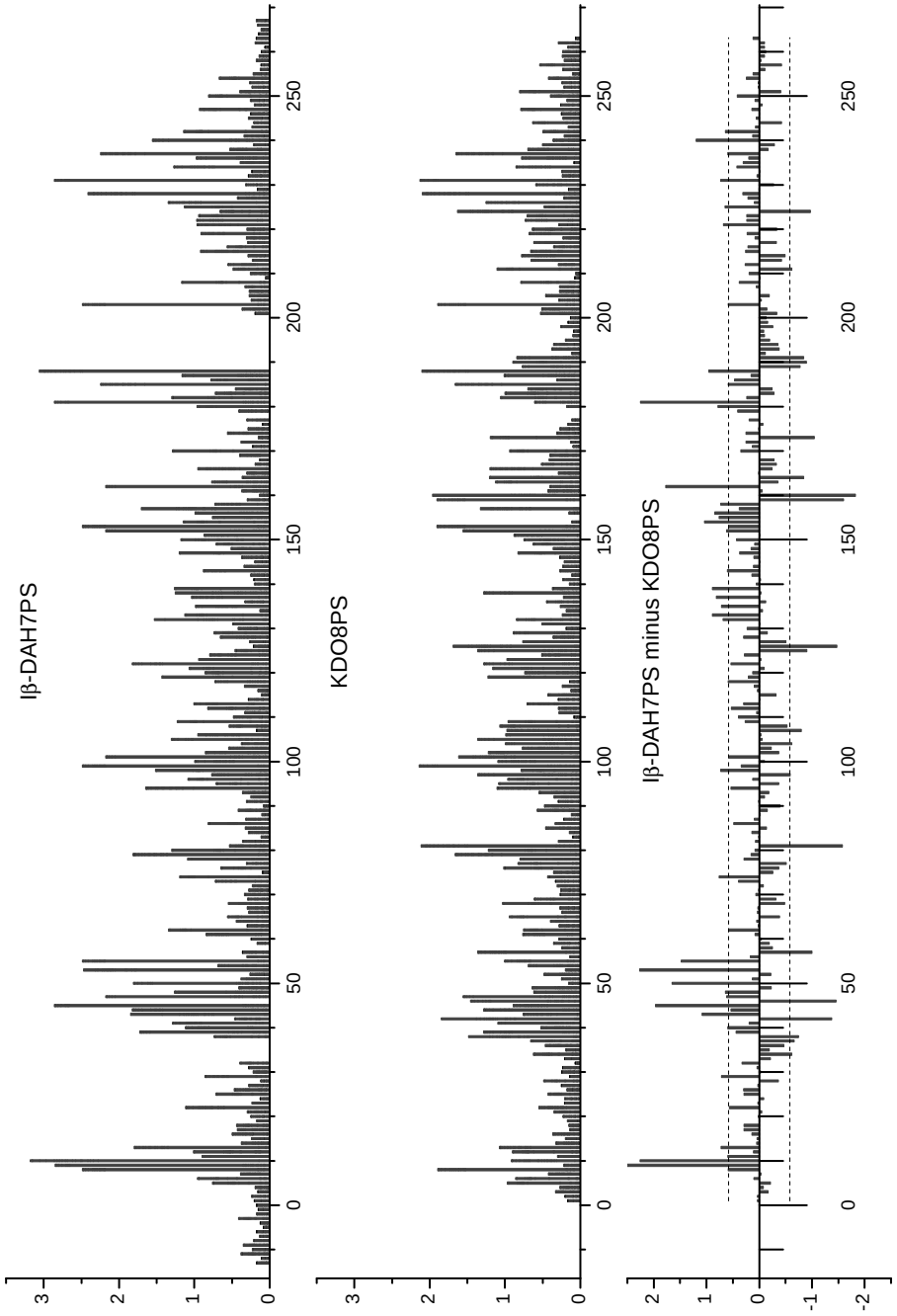


Figure 2.1: Relative entropy at each position of the profile HMMs for Iβ DAH7PS and KDO8PS and the calculated difference. The dashed lines are drawn at 1σ.

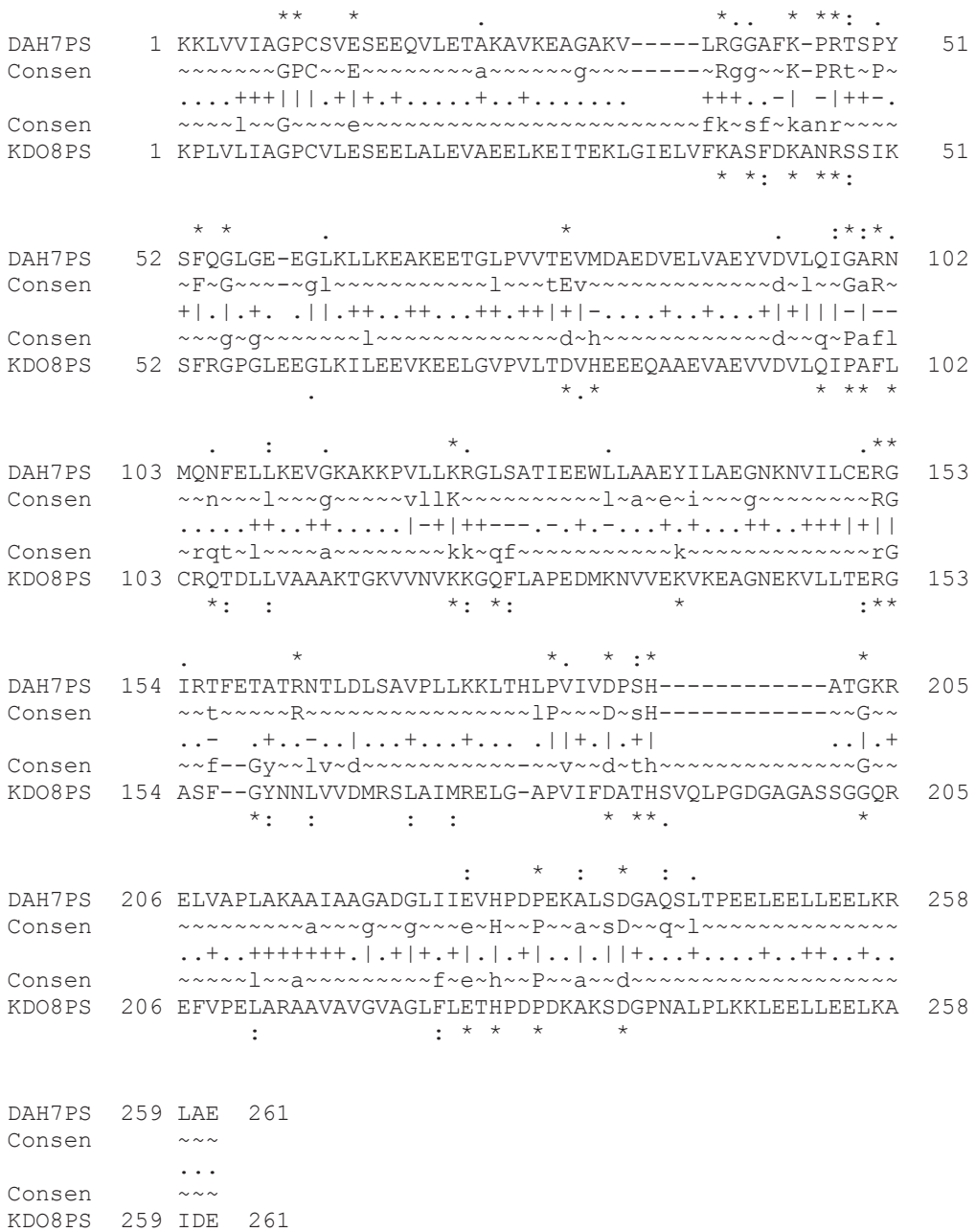


Figure 2.2: Core probabilistic model sequence for Iβ DAH7PS and KDO8PS aligned and annotated with conservation information from the original MSA. The profile comparison denotes similarity with (column score): = very bad match (< -1.5), - bad match (-1.5 to -0.5), . neutral match (-0.5 to 0.5), + good match (0.5 to 1.5,), | very good match (>1.5). The consensus sequence uses upper-case letters for amino acids that occur with $\geq$60 percent probability and lower-case letters for amino acids that have $\geq$40 percent probability. For non-conserved positions a tilde is used. The annotated conservation from the MSA uses standard Clustal codes.

to KDO8PSs is provided in Figure A.1.

2.2.3 Identified sequence similarities and differences

Using the original MSAs, profile HMM analyses, and structural interpretation, many similarities and differences were identified between Iβ DAH7PS and KDO8PS (Table 2.1). Positions were investigated if annotated as conserved in the MSA or profile HMM of either enzyme, or if the difference in relative entropy between the enzymes was greater than 1σ.

Table 2.1: Sequence features in Iβ DAH7PS and KDO8PS. Upper-case letters denote conservation while lower-case letters indicate variation.

Position	DAH7PS		KDO8PS	Structural context
8	G	=	G	Close packing; little side-chain room
9	P		p/m	
10	C		c/n	Metal ligand; Asn in metal-independent KDO8PS
12	V	≈	i/v/l/m	
13	E	=	E	H-bonds to loop with D237 via H228 and main-chain amide of 235
33–37	–		α1β2	α1β2 loop extension absent in Iβ DAH7PS
38	f/l	=	F	Buried hydrophobic
39	R		K	PEP binding site (carboxylate), H-bond to E226
41	G		S	S41 H-bonds K39, D79, main-chain carbonyl of 10
42	a		f/y	Buried hydrophobic; similar to f/y43
43	f/y		d	f/y43 stacks at interface area; D43 H-bonds to R54
44	K	=	K	PEP/A5P/E4P binding
45	–		A	
46	P		N	N46: A5P binding
47	R	=	R	A5P/E4P binding
48	s/t	=	s/t	A5P/E4P binding
49	s	=	s	

continued on next page

Table 2.1 – continued from previous page

Position	DAH7PS		KDO8PS	Structural context
50	P			P50 stacks at interface area
51	Y			Y51 stacks at interface area
53	F	=	y/f	
54	Q		R	Q54 H-bonds with main-chain of 51 and 52 and with Y138; R54 H-bonds D43 and D107 from adjacent subunit
55	G	=	G	
57	g	=	G	
62	L	=	l/m	Buried hydrophobic
65	L	=	l/f	Buried hydrophobic
76	i/v	=	i/v/l	
77	i/v	=	i/v/l	
78	t	=	t	
79	E		D	PEP binding site (nucleophilic water)
80	v	=	v	Buried hydrophobic
81	l/m/v		H	l/m/v81 buttresses K44 and P46; H81 H-bonds to main-chain of D43
85	d		e/q	H-bonds to main-chain of 81
94	D	=	D	H-bonds to residues on adjacent helices
96	l	=	l	Buried hydrophobic
97	Q	=	Q	PEP binding site (carboxylate), H-bonds to nucleophilic water
98	i/v	=	i/v/l	Buried hydrophobic
99	G		P	Limited side-chain space
100	a	=	A	Limited side-chain space
101	R		F	R101 E4P and PEP binding sites (phosphates); F101 stacks with F126 from adjacent subunit
102	n		L	N102 H-bonds with N105 from same and adjacent subunit; L102 from same and adjacent subunit form hydrophobic pair
104	Q		R	Q104 buttresses R101; R104 A5P binding (phosphate) in adjacent subunit

continued on next page

Table 2.1 – continued from previous page

Position	DAH7PS		KDO8PS	Structural context
105	N		Q	N105 H-bond pair over interface; Q105 H-bond to waters, pairing stymied by L102
106	F		T	F106 buttress main-chain of R47, stacks with Y138; T106 H-bonds with K138
107			D	D107 H-bonds with R54 from adjacent subunit
108	L	=	L	Buried hydrophobic
109	L	=	v/l/i	Buried hydrophobic
113	G	=	a	
119	V	=	v/i	Buried hydrophobic
120	L		N	L120 hydrophobically spacefilling; N120 weak interaction with Q97
121	L	=	v/i	Buried hydrophobic
122	K	=	K	PEP binding site (carboxylate/phosphate), H-bonds with D185
123	R		K	Buried, both H-bond to E151, R123 H-bonds also to D166
124	g	=	g	
125	m/l		Q	m/l125 space fill; Q125 H-bonds to R152
126	s/a		F	F126 stacks with F101
131	E		d/e/q	E131 makes H-bonds to R155 of adjacent subunit; d/e/q does the same with N162
137	e			
138	Y		K	Y138 stacks with F106 and H-bonds with Q54; K138 H-bonds with T106 and main-chain of R47 in adjacent subunit
139	I			Buried hydrophobic
143	G			
144	N			N144 H-bonds with main-chain of 117 and 146

continued on next page

Table 2.1 – continued from previous page

Position	DAH7PS		KDO8PS	Structural context
150	C	=	c/t	
151	E	=	E	H-bond to 123 and main-chain of 153
152	R	=	R	PEP binding site (phosphate)
153	G	=	G	Main-chain H-bonds to 151, limited side-chain space
154	I	=		Buried hydrophobic
155	R		–	H-bonds with E131 of adjacent subunit
156	t	=	c/s/t/n	On core tetramer loop
157	f/y	=	F	On core tetramer loop, next to equivalent residue of adjacent subunit
158	e/d		–	H-bonds to R123 in adjacent subunit
159			G	On core tetramer loop
161			n/d	H-bonds to main-chain of 193 on β7α7 loop
160	a/y		Y	On core tetramer loop, hydrophobic space, H-bonds to D166 of adjacent subunit
162	R		N	R162 H-bonds with D237 (metal ligand) and Q240; N162 potential H-bonds with adjacent subunit r168
163	N		L	
164	T		V	
166	D	=	D	H-bonds with R123 in Iβ DAH7PS, in KDO8PS interacts with 168
168	s/t		r/l	r/l interaction with D166
181	P	=	P	
182	i/v	=	V	Buried hydrophobic
185	D	=	D	Metal-binding site, H-bonds to H188 and K122
186	p/v	≈	a/v	
187	s/t	=	T	187 H-bonds to main-chain of 226
188	H	=	H	Metal ligand, H-bonds D185
189–200	–		β7α7	β7α7 loop absent in Iβ DAH7PS, Q191 A5P binding
201			s	S201 A5P binding (phosphate)
203	G	=	G	

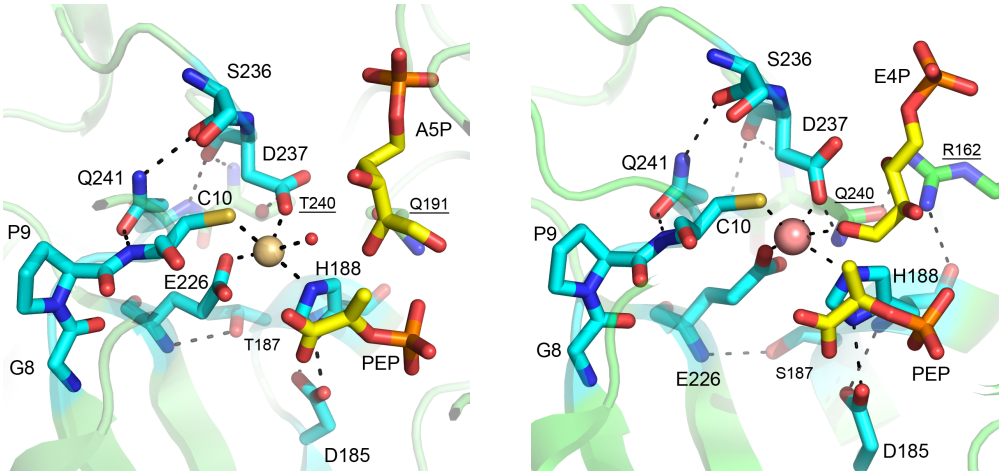
continued on next page

Table 2.1 – continued from previous page

Position	DAH7PS		KDO8PS	Structural context
211	l	=	L	Buried hydrophobic
214	A	=	a/s	Buried hydrophobic
224	m		F	PEP binding site (carboxylate)
226	E	=	E	Metal ligand, H-bond to K/R39
227	V			Buried hydrophobic
228	H	=	H	H-bonds to E13 and main-chain of 230
231	P	=	P	
234	A	=	A	
236	s/c		s/c	If metal-dependent and C, potential to form disulfide with C10
237	D	=	D	Metal ligand, A5P binding
238	g	=	g	
240	Q		n/s/t	Q240 H-bonds to D237, E226 and R162
241	s/q		m/a	q241 stabilises conformation of the β8α8 loop

2.2.4 Structural context of conserved residues in Iβ DAH7PSs and KDO8PSs

Using crystal structures of both a Iβ DAH7PS (*Pfu*DAH7PS) and KDO8PS (*Aae*KDO8PS) the sequence features listed in Table 2.1 were grouped into several areas. *Pfu*DAH7PS was chosen as it is an example of an Iβ DAH7PS with no (feedback-inhibition associated) decorations to the core TIM-barrel construction, and therefore the structure of *Pfu*DAH7PS is very similar to that of KDO8PS. *Aae*KDO8PS is one of the best characterised KDO8PSs, the crystal structure of which contains no disordered regions. Below, the similarities and differences in each group of conserved residues are given a structural context. This structural interpretation is important to make sense of whether what has been highlighted purely from comparison of sequence, may be important for determining substrate selection in a direct or potentially indirect manner.



(a) **(b)**

Figure 2.3: Metal-binding region of (a) *Aae*KDO8PS (PDB code 1FWW) and (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). The carbon atoms of residues primarily involved in metal binding are coloured cyan while those of PEP and A5P/E4P are coloured yellow. Positions that are conserved but differ in amino acid between the two enzymes are underlined.

For many residues that were identified as being conserved in both enzymes, on interpretation of crystal structures it was apparent that they were often buried hydrophobic residues, and their function was purely structural as part of the TIM-barrel fold. These residues will not be further discussed.

Metal-binding region

The metal-binding regions of Iβ DAH7PS and KDO8PS are each highly conserved and also very similar between the two enzymes (Figure 2.3). In this region, nearly every contributing residue is conserved across both enzymes, consistent with the common requirements: the ability to bind a divalent metal ion and to form a (partial) binding site for two substrates, PEP and a phosphorylated aldose sugar.

There are four metal ligands: C10, H188, E226 and D237. These are absolutely conserved in all sequences with the exception of C10, which is replaced by Asn in the case of metal-independent KDO8PSs. The other

non-metal-ligand conserved residues in this area seem to be important for precisely orienting loops and residues that are directly involved with metal binding and lining this area of the active site. D185 seems to play a role positioning metal-ligand H188, hydrogen bonding with its main chain and side chain. While S236, Q241, G8 and P9 are all involved in loop positioning to facilitate correct placement for (C/N)10 and D237. The residue at position 240 also has hydrogen-bonding interactions with D237. In KDO8PS, position 240 is either Asn, Ser or Thr (Figure 2.3a), whereas in Iβ DAH7PS it is conserved as Gln, which additionally hydrogen bonds to a conserved Arg at position 162 (Figure 2.3b). This difference is further discussed below.

In *Aae*KDO8PS, the crystal structure of which is used in Figure 2.3a, position 241 is Gln. This is uncommon in KDO8PS, where the residue at this position is more likely to be either Met or Ala. It is however very similar to Iβ DAH7PS, where the position is occupied by either Ser or Gln. In other KDO8PS structures the loop bearing residue 241 is sometimes disordered, but in the engineered metal-dependent mutants of *Nme*KDO8PS where the loop is partially ordered (PDB code 3FYO[39]), it is Ala and there are no substituting interactions.

PEP-binding pocket

Like the metal-binding region, the binding pocket for PEP is also well conserved within and between Iβ DAH7PS and KDO8PS. It is partially formed by the residues that contribute to the metal-binding region, but also by conserved residues and water molecules (Figure 2.4).

The phosphate moiety of PEP is hydrogen bonded to R152, the amide of A100, and to the amine of K122. In Iβ DAH7PS there is an extra hydrogen-bonding interaction with R101 (F101 in KDO8PS), however it is expected that with both substrates bound, R101 may also interact with the phosphate of E4P.

The Lys residue at position 122 hydrogen bonds with D185, which

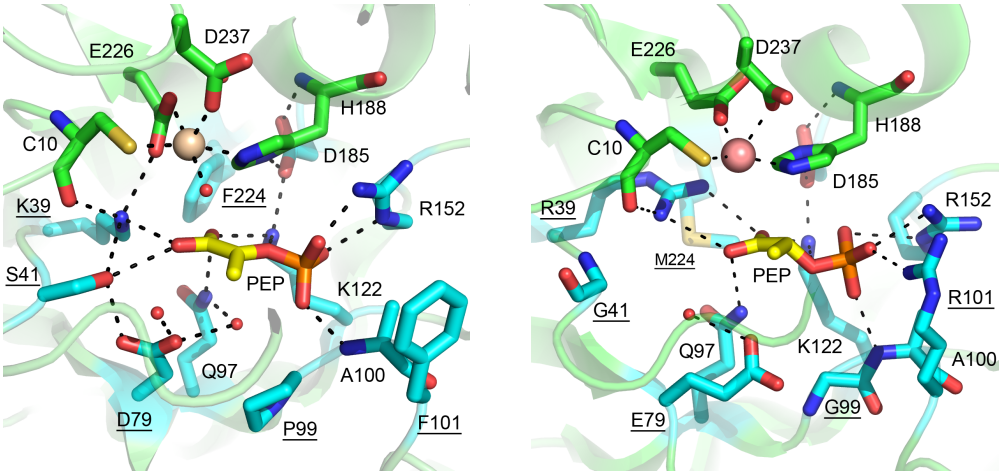


Figure 2.4: PEP-binding region of (a) *Aae*KDO8PS (PDB code 1FWW) and (b) *Pfu*DAH7PS (PDB code 1ZCO). The carbon atoms of metal ligands are coloured green, while those of all other conserved residues are coloured cyan and of PEP yellow, while Cd^{2+} is wheat, and Mn^{2+} is rose. Residues conserved, but different in each protein, are underlined.

has interactions with the metal-ligand H188. K122 also interacts with PEP through the phosphate and/or carboxylate moieties. In both proteins a conserved Gln at position 97 also interacts with the PEP carboxylate moiety.

In KDO8PS the residue D79 has hydrogen-bonding interactions with two water molecules, one of which is on the *re* face of PEP, and is also hydrogen bonded to Q97 (Figure 2.4a). The other water is hydrogen bonded to K44 (not shown). In the crystal structure of *Pfu*DAH7PS (Figure 2.4b) the former water molecule is not modelled, but the latter is observed. Additionally, in Iβ DAH7PS position 79 is a conserved Glu rather than Asp, which although offers the same functionality, the side chain is one carbon longer and sits in a slightly different position (Figure 2.4). The reason for this conserved difference is not obvious.

The residue at position 39 is also a conserved difference between Iβ DAH7PS and KDO8PS. This position is always Arg in Iβ DAH7PS and Lys in KDO8PS. In KDO8PS K39 hydrogen bonds with E226 to position this

metal ligand, and also to the main-chain carbonyl of residue 10, to residue S41, and to the PEP carboxylate. In Iβ DAH7PS, R39 has interactions with only the PEP carboxylate and the main-chain carbonyl of residue 10, and it is not obvious why KDO8PS requires a Lys, or whether some of the roles the Lys plays are substituted for in other ways in Iβ DAH7PS. However, it would seem that the combination of residues at positions 39, 41 and 79 in both proteins are correlated.

The only other difference in the PEP-binding pocket is at position 224, which in KDO8PS is a conserved Phe, and in Iβ DAH7PS is a moderately conserved Met. The residues seem to have only a space filling role and again it is not obvious why Phe is conserved in KDO8PS.

β7α7 loop extension

KDO8PSs possess an extended β7α7 loop that is 12 residues longer than in Iβ DAH7PS (Figure 2.5). While most of the residues that compose the loop extension (between residues 188 and 205) are variable, position 191 is conserved as Gln. Q191 interacts with C3-OH of A5P (or C4-OH depending on how A5P is modelled) and the conserved PEP-phosphate binding residue R152 (Figure 2.5a), and thus would appear to have some role in substrate selection. The C3-OH also interacts with D237, and for the equivalent hydroxyl (C3-OH) in Iβ DAH7PS this is the sole interaction (Figure 2.5b). Immediately subsequent to the loop extension in KDO8PS is a conserved serine residue at position 201, which forms part of the binding site for the phosphate moiety of A5P. The main-chain carbonyl of position 193 in the loop extension has hydrogen-bonding interactions with the side chain of the residue at position 161 of the adjacent β6α6 loop (which tends to be either Asn or Asp).

In Iβ DAH7PS the conserved residue R162 occupies a similar space to Q191 of the loop extension, and interacts with a conserved Gln at position 240 and with D237. In KDO8PS, the equivalent residue at position 162 is a conserved Asn, which interacts across the subunit interface (*vide infra*), while at position 240, the residue tends to be either Asn, Ser or Thr, which

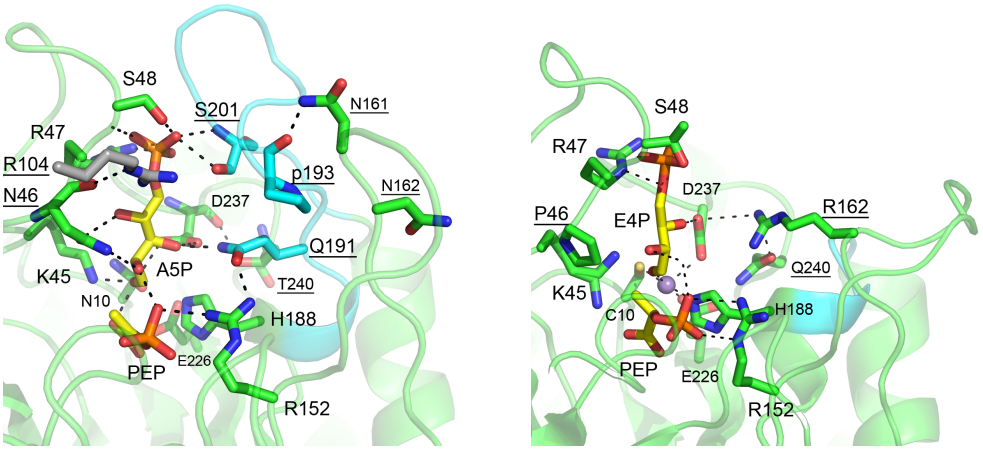


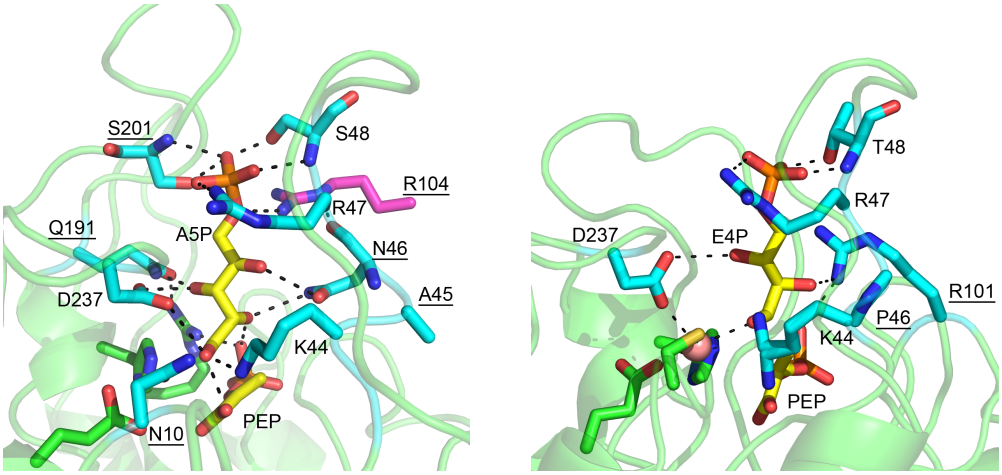
Figure 2.5: β7α7 loop of (a) *Aae*KDO8PS (PDB code 2NX3) and (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). The β7α7 loop is coloured cyan and the carbon atoms of PEP and A5P yellow. In (a) the carbon atoms of the interdigitating R104 from the adjacent subunit are coloured grey. Residues that are conserved but differ between the two proteins are underlined. Non-conserved residue letter codes are in lower case.

interacts with D237.

KDO8PSs also have an extended α1β2 loop compared to DAH7PSs. This five-residue extension at the opposite end of the barrel from the active site does not have any important interactions, is of variable composition, contains no conserved residues, and it seems is of little significance.

A5P/E4P-binding pocket and motifs

The binding region for A5P and E4P is formed by the previously discussed areas, as well as a conserved KANR(S/T) (in KDO8PS) or KPR(T/S) (in Iβ DAH7PS) motif that runs the length of the sugar. The motif in KDO-8PS provides hydrogen-bonding interactions with the hydroxyl groups that decorate A5P, as well as to the phosphate moiety positioned at the opposite end of the active site from PEP and aldehyde of A5P. The equivalent motif in Iβ DAH7PS does not provide any direct interactions with the hydroxyl groups



(a) **(b)**

Figure 2.6: The phosphorylated aldose binding region of (a) *Aae*KDO8PS (PDB code 2NX3) and (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). The carbon atoms of conserved residues are coloured cyan, while those of residues contributed from the adjacent subunit are coloured pink, and of PEP and the phosphorylated aldoses are coloured yellow. Residues conserved but different between the two proteins are underlined.

on E4P, but provides a Lys in a similar position proximal to the aldehyde of E4P, and residues to bind the phosphate moiety.

In KDO8PS, the Lys of the ^{44}KANR(S/T)^{48} motif has hydrogen bonds with the A5P carbonyl, and the Arg and Ser/Thr interact with the phosphate moiety of A5P. N46 hydrogen bonds with both C2-OH and C4-OH. Interdigitating from the adjacent subunit is R104, which has interactions with the phosphate moiety, and the main chain of N46.

Other interactions with A5P are through Q191 with C3-OH (as previously discussed), and D237 with C3-OH and the A5P carbonyl. The carbonyl of A5P also interacts with N10, in this metal-independent structure, and presumably interacts with the metal ion in the metal-dependent enzymes, similar to E4P in Iβ DAH7PS. Also as previously mentioned, S201 interacts with the phosphate moiety.

In Iβ DAH7PS the conserved motif is ^{44}KPR(S/T)^{48}, which is one

residue shorter, but the Lys, Arg and Thr have the same roles as in KDO8PS (Figure 2.6b). In most structures of Iβ DAH7PS the Pro of the motif is modelled with the main-chain carbonyl pointing toward E4P, with modelling studies indicating a potential interaction with the C2-OH.[87]

As for KDO8PS, in Iβ DAH7PS it is predicted that D237 will interact with C3-OH of E4P, and the carbonyl of E4P will coordinate to the metal ion. It is imagined that when E4P (and PEP) are bound (rather than E4P just modelled) in Iβ DAH7PS the side-chain position of R101 may move to mimic the role of the subunit-interdigitating R104 of KDO8PS.

Active-site subunit interface

The quaternary structure of both Iβ DAH7PS and KDO8PS is tetrameric, and assembly of the monomeric TIM-barrels results in the active site of each monomer being packed close to the adjacent subunit. Unsurprisingly, the residues that contribute to the interface area between two subunits, which are in close proximity to an active site, are conserved. However, the identity of the conserved interface residues are different between Iβ DAH7PS and KDO8PS. For each dimer (of the tetramer, defined by the interface closest to the active site) the interface is reciprocally mirrored, such that subunit B interfaces with the active site of subunit A, and subunit A interfaces with the active site of subunit B. Hence the whole dimer interface when analysed can be reduced to looking at interface interactions near one active site.

In KDO8PS the interface can be split into areas. Firstly, a hydrophobic region, where F101 and L102 from one subunit pack, respectively, against F126 and L102 of the adjacent subunit (Figure 2.7a). The second area has more polar interactions. From the adjacent subunit R104 interdigitates to become one of the ligands of the phosphate moiety of A5P, and has interactions with the main-chain carbonyl of N46 in the adjacent subunit. These residues (R104, F101 and L102) are part of a conserved ^{99}PAFLxR^{104} sequence motif in KDO8PS. The main-chain carbonyl of R104 interacts with the amine of K138 from the same subunit. K138 also hydrogen bonds with T106, and to the

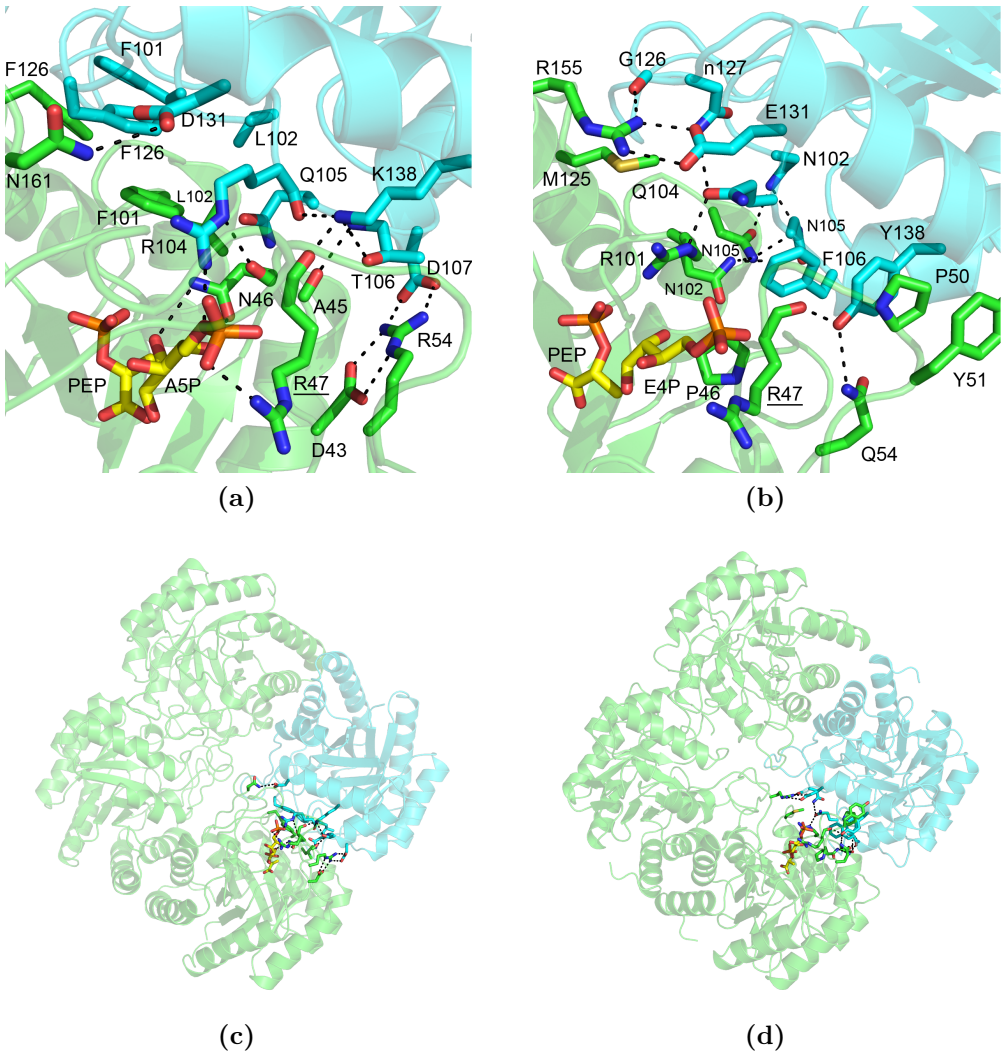


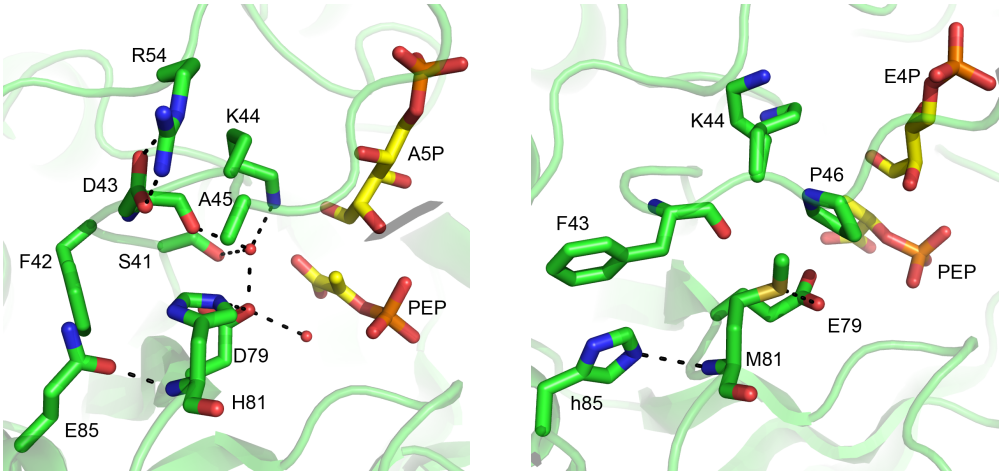
Figure 2.7: The subunit interface adjacent to the active sites of *Aae*KDO8PS (PDB code 2NX3) (a) and (c), and *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]) (b) and (d). The carbon atoms of PEP and A5P/E4P are coloured yellow, residues that belong to the same subunit as the pictured active site green, and those from the adjacent subunit cyan. Residues that are conserved and the same in both proteins are underlined.

main-chain carbonyl of R47 in the adjacent subunit. The interface is further strengthened by interactions with R54 (held in place by D43) across the interface to D107. The net effect seems to be that this interface in KDO8PS is very important in constructing the A5P binding region of the active site. The KANR(S/T) motif is held firmly in place by residues from the adjacent subunit, which also contributes a residue to complete the A5P binding site.

In Iβ DAH7PS on the other hand, the relationship between the adjacent subunit and the equivalent E4P binding residues [KPR(T/S) motif] appears less intimate, and more hydrophobic in nature. Y51 and P50 from one subunit stack with Y138 and F106 from the adjacent subunit. Y138 also has hydrogen-bonding interactions across the interface with the main-chain carbonyl of R47 and with Q54. While the interface is still supporting the portion of the β2α2 loop that forms the E4P binding site, it is doing so less intimately than in KDO8PS. Where in KDO8PS R104 interdigitates across the interface to contribute to the phosphate binding area of the A5P binding site, in Iβ DAH7PS it is replaced by a Gln residue, which appears to buttress R101 into providing the same functionality. Whereas in KDO8PS the conserved sequence motif in this area is PAFLxR, in Iβ DAH7PS the equivalent motif, which is also conserved, is ^{99}GARNxQ^{104}. The hydrophobic interface region in KDO8PS is in Iβ DAH7PS replaced by complementary hydrogen-bonding interactions between N102 and N105 and their counterparts from the adjacent subunit.

Residues of structural importance

Many of the conserved residues identified in this analysis seem to be important for structural reasons, rather than being of any direct functional importance. In many cases, the residues have hydrophobic side chains and sit buried within the core of the TIM-barrel structure between β-sheets and α-helices. However, many other identified residues are on the periphery of the active site and may have an influence on the function of the enzyme and in some cases there are conserved differences in these residues between Iβ DAH7PS and KDO8PS.



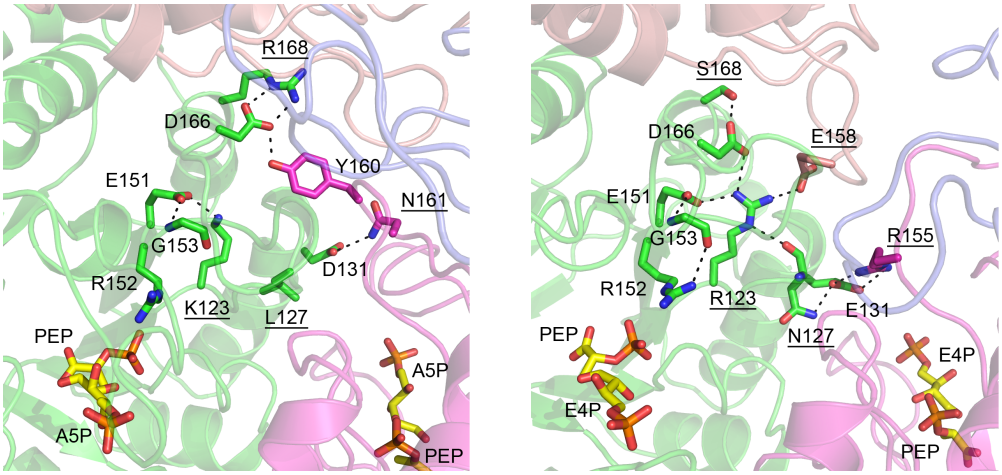
(a) **(b)**

Figure 2.8: Residues moderately conserved close to the conserved residues of the β2α2 loop and residue at position 79. (a) *Aae*KDO8PS (PDB code 2NX3) and (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). The carbon atoms of PEP and A5P/E4P are coloured yellow.

In KDO8PS a Glu or Gln residue at position 85 hydrogen bonds with the main-chain amide of the nearby residue at position 81, which is a conserved His (Figure 2.8a). This His residue has hydrogen-bonding interactions with the carboxylate moiety of conserved D79, which is central to a network of water molecules that sit between itself, PEP, D43, S41 and K44.

In Iβ DAH7PS position 85 is often an Asp (although in *Pfu*DAH7PS is a His), and would seem to have the same interaction with the main-chain amide of position 81 (Figure 2.8b). However position 81 is less conserved in Iβ DAH7PS than in KDO8PS, and varies between being Met, Leu or Val. The residue at this position seems to buttress E79 and the the β2α2 loop bearing K44 and P46. Although absent in the crystal structure of *Pfu*DAH7PS used in Figure 2.8b, in Iβ DAH7PS the same water molecules shown in Figure 2.8a are thought to be present, and therefore the same network exists that connects, in this case E79 with K44.[74]

In both Iβ DAH7PS and KDO8PS the phosphate moiety of the aldose sugar interacts with a conserved Arg at position 152 (Figure 2.9). In close



(a) **(b)**

Figure 2.9: Loop interactions near subunit boundaries in (a) *Aae*KDO8PS (PDB code 2NX3) and (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). Each subunit is coloured either green, pink, blue or rust. In KDO8PS, all residues are conserved with the exception of R168 and L127. In Iβ DAH7PS all residues are conserved except for N127. Residues that differ between the two proteins are underlined.

proximity to this residue is the interface region between the four subunits that compose the tetramer. Some of the residues in this area are conserved either as the same amino acid in both proteins, or as different amino acids, or are absent in one or the other.

Common to both proteins, apart from R152, are E151, G153 and D166, with both E151 and G153 seeming to be important for correctly placing R152. The residue at position 131 is also highly conserved as either Asp or Glu. The key difference between the two proteins centres on the conserved residue at position 123, which in KDO8PS is Lys, and in Iβ DAH7PS Arg. In KDO8PS, K123 hydrogen bonds with E151, which in turn interacts with the main-chain amide of G153, a relationship maintained in Iβ DAH7PS despite the interchange of the Arg for Lys. However, the longer and more richly functionalised Arg affords additional interactions, to the main-chain carbonyl of N127, to the carboxylate of D166, but also to E158 from an adjacent subunit. In both proteins, the conserved residue D166 interacts with a residue

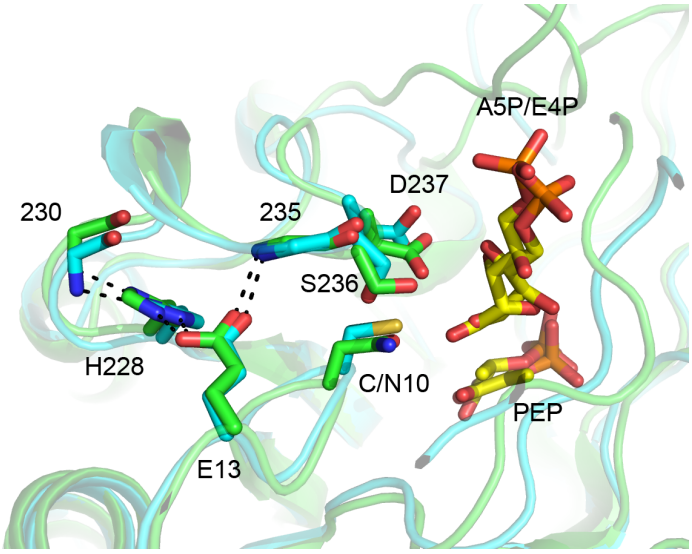


Figure 2.10: Superposition of *Aae*KDO8PS (coloured green) and *Pfu*DAH7PS (coloured cyan) shows the conserved residues responsible for positioning the β8α8 loop. The carbon atoms of PEP and A5P/E4P are coloured yellow. Only the main chain of residues 230 and 235 is shown, and both residues are not conserved. The RMSD for the superposition of Cα atoms for the subunits shown is 1.074 Å.

at position 168, which is often an Arg (but also sometimes Leu) in KDO8PS and Ser or Thr in Iβ DAH7PS.

There are also differences in the interface interactions between the two proteins. In KDO8PS Y160 hydrogen bonds to D166 of an adjacent subunit, and D131 hydrogen bonds to N161 of the same subunit. In Iβ DAH7PS the pattern of inter-subunit interactions is different. Here, E131 interacts with R155 from an adjacent subunit, but also R123 hydrogen bonds with E158 from a different adjacent subunit. In KDO8PS, positions 155 and 158 (respectively conserved as R and D/E in Iβ DAH7PS) are absent, and conformations of the adjacent subunits' loops are different. This region is distal from the active sites in both proteins (≈15 Å) and it is difficult to understand the consequences of the different loop conformation and interactions.

Residues E13 and H228 were identified as being absolutely conserved in both Iβ DAH7PS and KDO8PS. E13 hydrogen bonds with H228 and the main-chain amide of position 235 and H228 also hydrogen bonds with the

Table 2.2: Sequence features in metal-dependent and metal-independent KDO8PS. Upper-case letters denote conservation while lower-case letters indicate variation.

Position	Metal dependent		Metal independent	Structural context
9	p		i/m	
10	C		N	Metal-ligand Cys; Asn in metal independent
72	l		f/y	Buried hydrophobic
114	G		p/a/g	
150	t/c		C	Of (C/T)ERG motif
181	P	=	P	Only difference in relative entropy
190	s	≈	s/a	
236	s/c	≈	c	In proximity to metal-ligand Cys10

main-chain amide of residue 230 (Figure 2.10). The role of these residues seems to be to hold the β8α8 loop in a twisted conformation, such that S236 and D237 are positioned pointing in toward the active site. If the loop was not constrained by these residues then these two residues, intimately involved in metal and substrate binding, would presumably be incorrectly positioned.

2.3 Comparison of the two forms of KDO8PS

The pool of KDO8PS sequences was split into two sequence sets, being either metal dependent (630 sequences) or independent (470 sequences) based on the presence of the metal-dependency defining Cys or Asn. The sequences in both sets were realigned, and profile HMMs were created. The profile HMMs were aligned with each other and the core probabilistic model sequence was annotated with the conservation indications from the MSA (Figure 2.11). The relative entropy at each position was calculated, as was the difference in relative entropy at each position (Figure 2.12). Several differences in conservation were identified using these analyses and are listed in Table 2.2.

Consistent with the predicted recent evolution of the metal-independent KDO8PSs, the set of metal-independent sequences is less diverse than the metal-dependent set. The side effect of the high homology between the metal-independent sequences is a reduced relative entropy for positions which are highly conserved, compared to positions which are less well conserved. This makes it harder to identify conserved residues from the differences in magnitude of relative entropy at each position, and from the predicted conservation (computed from the magnitude of the probability scores). Complementing the profile HMM analysis with the MSA analysis helped to overcome this limitation.

2.3.1 Differences in sequence associated with metal dependency

The sequence analysis of the two forms of KDO8PS identified that the identity of the residue at position 10 (Cys or Asn) strongly correlates with metal dependency. The analysis indicated that the residue at position 9 also correlates with whether the adjacent residue is Cys or Asn. In metal-dependent KDO8PSs this residue is most likely to be Pro, whereas in metal-independent KDO8PSs the residue is likely to be either Ile or Met. It has been suggested that the more conformationally restrictive Pro helps to position the Cys in metal-dependent KDO8PSs to optimise metal binding (Figure 2.13).[24,68]

The identity of the residue at position 236 also had some correlation to metal dependency. In metal-dependent KDO8PSs this residue is very likely to be Ser, whereas in metal-independent forms it is more variable: sometimes Ser but often Cys. In studies of *Nme*KDO8PS where the metal dependency defining Asn was mutated to Cys, Cys246 (positionally equivalent to 236) was found to form an unwanted disulfide bond with Asn23Cys, which has been suggested as a reason why in metal-dependent KDO8PSs this residue is mostly likely a Ser.[39]

In both DAH7PS and KDO8PS there is a conserved (C/T)ERG motif

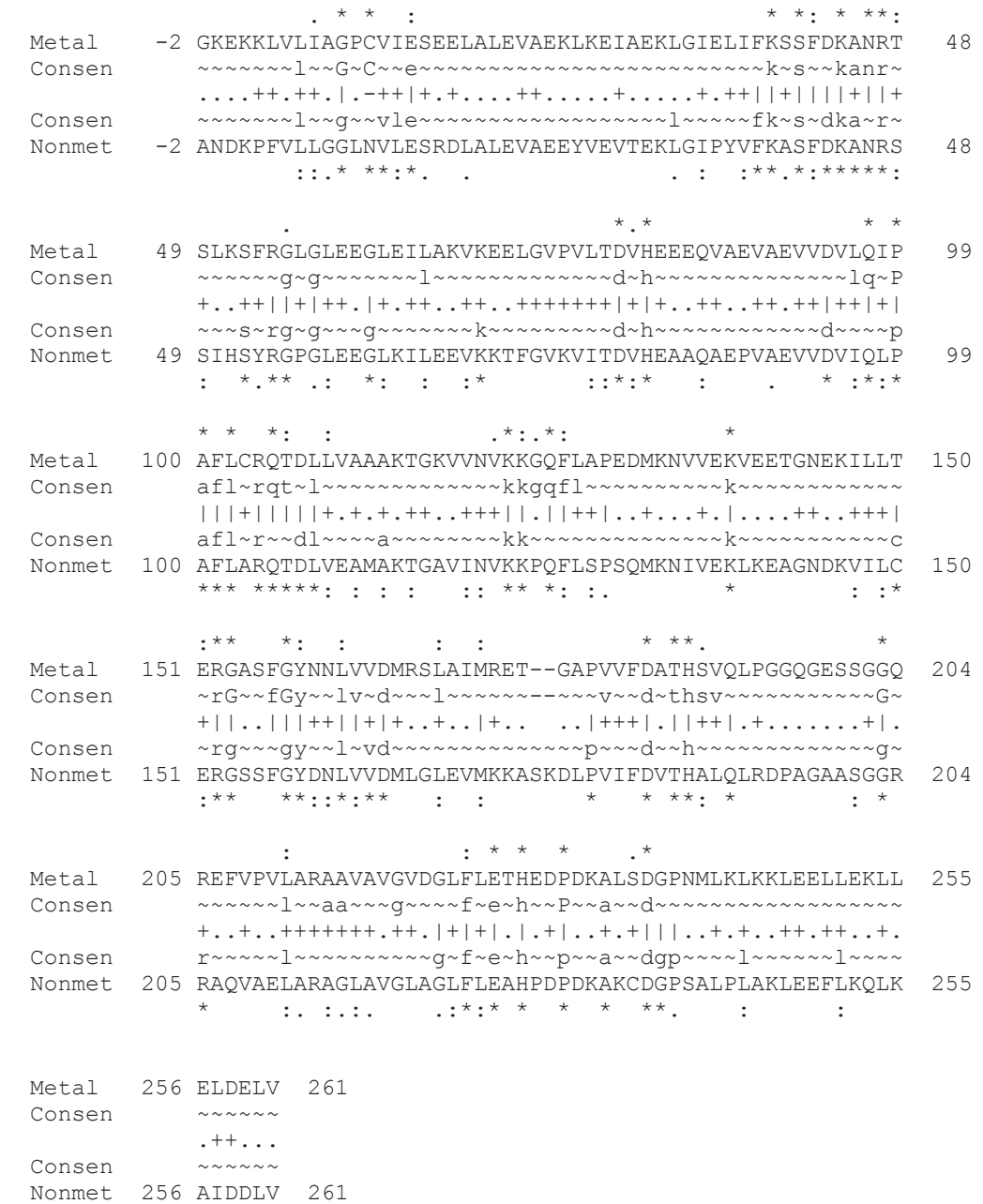


Figure 2.11: Core probabilistic model sequence for metal-dependent and metal-independent KDO8PSs aligned and annotated with conservation information from MSA. The profile comparison denotes similarity with (column score): = very bad match (< -1.5), - bad match (-1.5 to -0.5), . neutral match (-0.5 to 0.5), + good match (0.5 to 1.5,), | very good match (>1.5). The consensus sequence uses upper-case letters for amino acids that occur with $\geq$60 percent probability and lower-case letters for amino acids that have $\geq$40 percent probability. For non-conserved positions a tilde is used. The annotated conservation from the MSA uses standard Clustal codes.

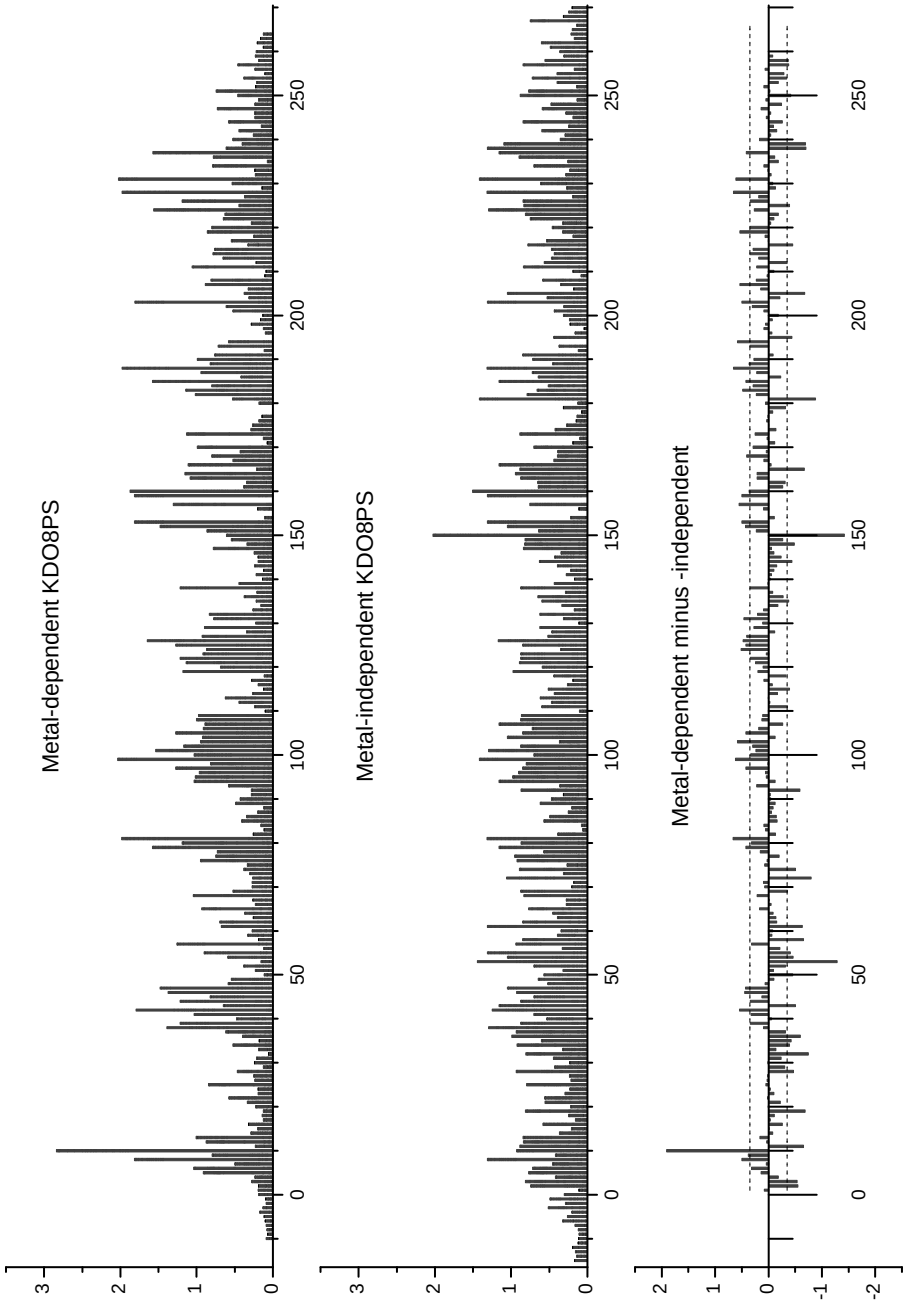


Figure 2.12: Relative entropy at each position of the profile HMMs for both forms of KDO8PS and the calculated difference. The dashed lines are drawn at 1σ.

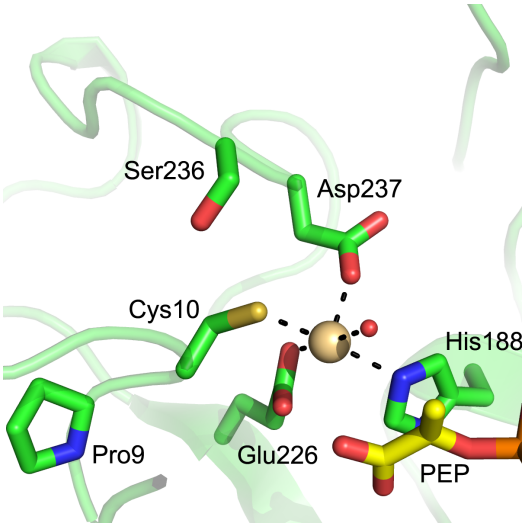


Figure 2.13: Structure of *Aae*KDO8PS (PDB code 1FWW) showing the metal-binding area of KDO8PS. The carbon atoms of PEP are coloured yellow and Cd^{2+} is coloured wheat.

on the β6 strand, as previously discussed, of which the Arg coordinates the PEP phosphate. In both enzymes the first residue of the motif is found as either a Cys or Thr, however in all metal-independent KDO8PSs the first residue is always Cys. In the metal-dependent *Aae*KDO8PS, the Thr (T150) hydrogen bonds to a conserved Asn (N110) on the β5 strand, suggesting that its role is to stabilise the TIM-barrel (Figure 2.14). Whether the residue is a Thr or Cys would not impact on this function, suggesting that the conservation in metal-independent KDO8PSs is merely significant of only the lack of divergence in metal-independent KDO8PSs. It also seems that given the distance of the motif from the metal-binding site, whether the first residue is Thr or Cys is unlikely to be important for metal dependency.

The other conservation differences identified in this analysis were either artefacts of the relative entropy calculations and therefore not real due to low sequence divergence, or were unlikely to be associated with metal dependency. Position 72 is on the end of helix α2 in a hydrophobic environment, and in both forms the identity of the residue is hydrophobic. Position 114 is on the β5α5 loop, in a tightly twisted region where both Gly or Pro would offer access for the main chain to the desired conformation. This loop region is on

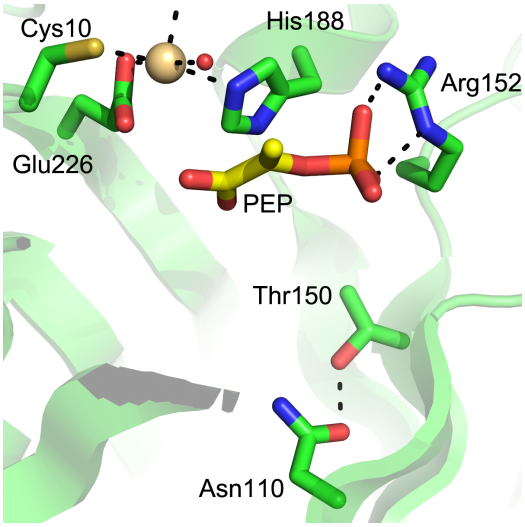


Figure 2.14: Structure of *Aae*KDO8PS (PDB code 1FWW) showing the interaction between Thr and Asn between adjacent β-sheets in the TIM-barrel core. The carbon atoms of PEP are coloured yellow and Cd^{2+} is coloured wheat.

the opposite side of PEP from the metal-binding site, and it is difficult to rationalise why Gly at this position may be conserved in metal-dependent KDO8PSs and how it could be related to metal dependency. Position 181 is identified as being potentially different from relative entropy calculations, however in both forms is almost always Pro. For position 190, which is on a short helical twist immediately before the beginning of the β7α7 loop, the Ser has no obviously interactions or role.

2.4 Discussion

Approximately 2000 DAH7PS sequences were assembled and of these around 375 were identified as being of type Iβ DAH7PS based on the recognition of sequence features that distinguish the three types. The Iβ DAH7PS sequences were analysed by MSA and profile HMM analyses, and these analyses were compared to the same treatment of approximately one thousand KDO8PS sequences. This identified many residues that are conserved in both enzymes, consistent with the homology and predicted evolutionary

relationship, including the TIM-barrel fold and commonality of substrates and metal binding.

Many differences in conserved residues between the two enzymes were also identified. From the structural analysis, many of these differences have direct interaction with the phosphorylated aldose substrate and give some clues as to what is responsible for the different substrate selection preferences of the two enzymes. The most obvious differences associated with substrate selection are:

- An extended β7α7 loop in KDO8PS
- KANR(S/T) motif (KDO8PS) versus KPR(S/T) motif (DAH7PS) on the β2α2 loop
- PAFLxR motif (KDO8PS) versus GARNxQ motif (DAH7PS) on the β4α4 loop

These enzyme-unique sequence features are investigated in some of the following chapters of this thesis.

Further to the analysis comparing Iβ DAH7PS to KDO8PS, the KDO-8PS sequences were split into two sets based on the presence of a Cys or Asn, which is the hallmark of metal dependency. The same type of analysis, MSA and profile HMM, was used to compare the two forms of KDO8PS to identify differences associated with the difference in metal dependency. This showed that the two forms of KDO8PS are very similar. Apart from the Cys to Asn switch which defines metal dependency, this analysis identified only two other significant sequence differences. The first is the identity of the residue immediately prior to the Cys/Asn, as being either a Met (metal independent) or Pro (metal dependent). The second difference is the presence of either Cys or Ser on the β8α8 loop. The presence of a Ser is correlated with an enzyme being metal dependent, and of a Cys with metal-independent forms of KDO8PS. Both of these differences between the forms have been previously identified and assessed for their role in metal dependency.[39,68]

2.4.1 Issues encountered with analyses

Although the sequence analysis presented in this chapter revealed key differences both between KDO8PS and DAH7PS and between the two forms of KDO8PS, some issues and limitations in the analyses were also identified.

Analysis of the metal-independent KDO8PS sequences highlighted that a sample of sequences with overall high similarity lowers the median relative entropy over a profile HMM, which makes highly conserved positions less distinguishable from those that are more variable. However, the HMMER software does appropriately weight groups of sequences that are very similar (for example those from different strains of the same species) to reduce conservation bias. The problem of low sequence variation is not confined to HMM analysis: MSA analysis can likewise inadvertently identify positions that are conserved but not necessarily important for function because of the dilution in the correlation between conservation and functional importance.

Note also should be made that the prediction of conservation in a profile HMM uses the probabilities of amino acids at each position. While this overcomes to an extent the noise created by sequence outliers that somewhat hampers MSA conservation identification (gaps are particularly disruptive to the typical assignment of conservation), it has to be interpreted differently from MSA-derived conservation. A strong probability of a particular amino acid at a position infers neither that there is total conservation or that all amino acids at a position are very similar. Hence, the conservation of residues derived from a profile HMM can be different to that from the MSA that was used to create the profile. Indeed, as mentioned above, it is also possible for low probabilities (caused by low sequence variation) of very highly conserved positions to not be identified as conserved by this probability threshold method. Additionally, analysing differences in total relative entropy at each position alone will not highlight underlying differences in amino acid choice, only differences in conservation at a position. It is therefore important to also include the identities of the conserved residues rather than just the differences in conservation.

2.5 Summary

A collection of Iβ DAH7PS sequences was analysed by the creation and use of a MSA and profile HMM. This was compared to the same analysis of many KDO8PS sequences. These two analyses were compared to identify the similarities and differences in sequence conservation between the two enzymes. The highlighted sequence features were scrutinised in the context of representative crystal structure models of each enzyme to understand the role of the residues and the reason for the conservation. Reassuringly, the MSA and profile HMM analyses highlighted previously identified distinguishing sequence differences between the two enzymes. The same analysis was also used to compare metal-dependent to metal-independent KDO8PSs. This demonstrated that the two forms of KDO8PS are very similar: only three sequence differences were discovered that correlated with metal dependency.

Chapter 3

The metal dependency of *Acidithiobacillus ferrooxidans* KDO8PS

3.1 Introduction

One of the most intriguing aspects of KDO8PS is the relationship between metal-dependent and metal-independent forms and the role of the residues associated with metal binding. The precise role of the metal ion is unclear, and how metal-free variants compensate for the metal loss, when the natural metal-dependent enzyme is completely inactive when depleted of its divalent metal ion, needs to be clarified.

Some insights into metal dependence have been gained from several studies that interconverted the two forms of KDO8PS, converting enzymes from being natively metal-dependent to become metal-independent and vice versa.[24,28,39,67,68] These studies all mutated what is the primary determinant of metal dependence in KDO8PS, a residue at the end of the β1 strand present as a metal-ligand Cys (metal-dependent) or Asn (metal-independent).

Together with the Cys-Asn substitution, some of the studies also mutated other residues, the identities of which correlate to each form of KDO8PS.

3.1.1 Conversions to a metal-dependent form

Three separate studies have reported the conversion of the metal-independent *Eco*KDO8PS into a metal-dependent enzyme with remarkably variable results.[28,67,68] In all of the studies of the engineered metal-dependent *Eco*-KDO8PSs, metal ions enhance the activity of the enzyme, rather than the enzyme being truly metal dependent. That is to say, the enzymes retain residual metal-independent activity—a feature absent in wild-type metal-dependent KDO8PSs.

In the study by Oliynyk et al.,[68] N26C *Eco*KDO8PS in the absence of Mn^{2+} retained 14 percent of wild-type activity. However, in the presence of Mn^{2+}, the enzyme retained 22 percent of wild-type activity, representing an activation by this metal ion of 160 percent. The identity of the residue immediately preceding the Asn/Cys is also a conserved difference between the two forms of KDO8PS: in metal-dependent KDO8PSs (and DAH7PS), the metal-binding Cys exists as part of a Pro-Cys pair, and in metal-independent KDO8PSs the equivalent pair is Met/Ile-Asn. Oliynyk et al. coupled N26C with the mutation M25P, which resulted in an enzyme with 29 and 57 percent activity of wild-type in the presence of EDTA and Mn^{2+} respectively, corresponding to a 196 percent activation by the metal ion. While the activities of both single and double mutant proteins were greater in the presence rather than absence of Mn^{2+}, the Mn^{2+}-activated enzymes were less active than wild-type *Eco*KDO8PS and did not resemble contemporary metal-dependent enzymes, due to the high catalytic activity in the absence of metal ion.

A second study also created the metal-binding N26C *Eco*KDO8PS.[28] EDTA treated N26C *Eco*KDO8PS retained 6 percent of wild-type activity and was activated equally by Mn^{2+} and Cd^{2+} to 31 percent of wild-type activity, representing a 516 percent activation by the metal ions. Although

the mutant enzyme retained less activity than in the aforementioned study, the level of metal activation in this study was greater.

In the third study,[67] N26C *Eco*KDO8PS was found to have very low activity, less than 2 percent that of wild-type. However, contrary to the results of the other studies, the mutant enzyme was found to be inhibited by metal ions, as was the wild-type enzyme. This study used an alternative thiobarbituric acid assay to detect product formation rather than the more common continuous assay that measures the loss of PEP. The thiobarbituric acid assay can be sensitive to the presence of other species in the assay mixture, and it is possible this may account for the anomalous results.

The only other metal-independent KDO8PS to be converted to become metal-dependent is *Nme*KDO8PS.[39] The single mutation of Asn for Cys (N23C) created an enzyme where the presence of a divalent metal ion (Mn^{2+} or Cd^{2+}) was activating, but which also retained metal-independent activity (in the absence of metal ions). One of the conserved differences between the forms of KDO8PS (apart from the Cys/Asn) is the composition of a short, four-member sequence motif, located on the β8α8 loop, which includes the metal-ligand Asp (Figure 3.1). In bacterial metal-independent KDO8PSs the motif is CDGP, but in metal-dependent forms it is commonly SDGP, and the first residue is always Ser. Based on these observations, the N23C mutation of *Nme*KDO8PS was also studied in combination with three other active-site mutations. The roles of the residues in the motif were assessed by mutating Cys246, Asp247 and Pro249.

A key finding of the *Nme*KDO8PS study was that the Cys of the motif can unfavourably form a disulfide bond with the metal-ligand Cys (Figure 3.1), which is introduced in place of Asn in the conversion to a metal-dependent form. Interestingly, some metal-dependent KDO8PSs do have a Cys at this position. Additionally, all type II DAH7PSs (except one[88]) have a Cys in this position and it is known that *Mtu*DAH7PS and *H. pylori* DAH7PS (*Hpy*DAH7PS) are sensitive to oxidation.[56,89] It is therefore likely that metal-dependent KDO8PSs with this Cys residue may also be prone to disulfide bond formation, which may explain the variable results in the studies of

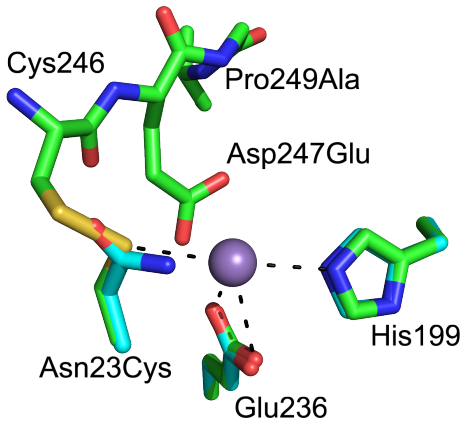


Figure 3.1: The metal-binding site of N23C/D247E/P249A *Nme*KDO8PS (PDB code 3FYO, carbon atoms coloured green) showing the partially occupied Mn^{2+} ion (coloured purple) and disulfide bond between Cys23 and Cys246, superimposed on the equivalent residues from wild-type *Nme*KDO8PS (PDB code 2QKF, carbon atoms coloured cyan). Unlike in the triple mutant, the β8α8 loop in wild-type *Nme*KDO8PS is disordered and not modelled in the crystal structure.

*Eco*KDO8PS.[39] In metal-independent KDO8PSs, it seems the Cys of the motif has a role in positioning the β8α8 loop, by hydrogen-bonding to the Asn of the β2α2 loop (Figure 3.1), which may compensate for the absence of a metal ion.[39] Unlike for *Eco*KDO8PS where the N26C mutated enzyme retained metal-independent activity, obligate metal-dependency was created in *Nme*KDO8PS when the Cys of the motif was mutated to Ser (C246S) in conjunction with N23C.

The other key finding from the *Nme*KDO8PS study is that the Asp of the motif is important for activity in both metal-dependent and metal-independent KDO8PSs. Moreover, the Asp was found to have a dual role as both a metal ligand and a provider of key interactions with A5P. In metal-dependent KDO8PSs, the metal ion secures the position of the β8α8 loop, but more specifically the Asp residue. As such, in the absence of the metal ion the β8α8 loop on which the Asp resides requires subtle modifications (such as the Pro residue immediately after Asp), evident in the differences to the Asp-containing sequence motif. Even with these modifications however,

in structures of *Eco*KDO8PS and *Nme*KDO8PS the β8α8 loop is relatively mobile. In *Eco*KDO8PS the side chain of Asp is in a different orientation than in *Aae*KDO8PS, and in *Nme*KDO8PS most of the β8α8 loop is disordered.

3.1.2 Conversions to a metal-independent form

Only for the very closely related KDO8PSs from the hyperthermophilic bacteria *A. aeolicus* and *A. pyrophilus* has a switch from a metal-dependent to a metal-independent form been accomplished.[24,28,67] In the conversion of *A. pyrophilus* KDO8PS (*Apy*KDO8PS), a metal-independent form was created (C11N) with 5–6 percent of wild-type activity.[28] Two studies have created metal-independent *Aae*KDO8PSs. A first created a metal-independent *Aae*-KDO8PS (C11N) that had 83 percent of wild-type activity.[67] A second study was more extensive and created a series of enzymes with multiple mutations (all of which included the requisite C11N mutation) in an attempt to recreate the evolutionary pathway between metal-dependent and metal-independent KDO8PSs (Figure 3.2).

In the latter study,[24] along with C11N, Pro10, Ser235 and Gln237 were mutated to Met, Pro and Ala respectively, mutations that represent conserved differences between the two forms of KDO8PS. In *Eco*KDO8PS Pro235 (*Aae*-KDO8PS numbering) was observed to bend the β8α8 loop to help position the following Asp residue (a metal ligand in metal-dependent KDO8PSs). The Pro, conserved in all metal-independent KDO8PSs, appears to improve the positioning of Asp in the absence of a metal ion. The series of mutant enzymes, C11N, P10M/C11N, C11N/S235P/Q237A, P10M/C11N/S235P/Q237A, all had metal-independent activity, and the activity (k_{cat}) decreased in order from the single mutant (46 percent of wild type) to the quadruple mutant (25 percent of wild type). Interestingly, Gln237, which has a role in securing the conformation of the β8α8 loop, is rare in KDO8PSs yet is conserved in Iβ DAH7PSs, and the only metal-independent KDO8PSs that have a Gln at this position are from plants.

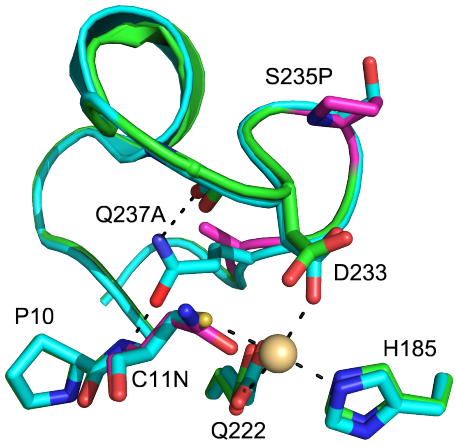


Figure 3.2: Superposition of wild-type (PDB code 1FWW, carbon atoms coloured cyan) and C11N/S235P/Q237A (PDB code 2NX3, carbon atoms coloured green) *Aae*KDO8PS. In both models the β8α8 loop is shown in cartoon representation. The carbon atoms of residues mutated in the triple mutant are coloured pink. The metal ion (coloured wheat) is from the wild-type structure. Q237 in the wild-type structure interacts with the main chain of residues 10 (β2α2 loop) and 232 (β8α8 loop).

3.1.3 *Acidithiobacillus ferrooxidans* KDO8PS

In this chapter a KDO8PS from the mesophilic bacterium *A. ferrooxidans* is examined. This KDO8PS was selected for analysis as it is predicted to be metal dependent, based on the presence of a putative metal-binding Cys rather than an Asn in the primary sequence. Moreover, *Afe*KDO8PS has near the C-terminus the SDGP motif (where the Asp is the putative metal-binding residue), the DGP portion of which is absolutely conserved among currently sequenced metal-independent KDO8PSs, but which is only sometimes present among the metal-dependent enzymes.[90] Being a mesophilic metal-dependent KDO8PS it is also a natural complement to the metal-independent *Nme*-KDO8PS, the other enzyme extensively studied in this thesis.

*Afe*KDO8PS was the subject of some previous characterisation and investigation in our laboratory: it had been cloned into an expression vector, a purification protocol developed, and some preliminary kinetic characterisation had been completed.[84] *Afe*KDO8PS was found to be metal dependent and

possess maximal activity at pH 7.2 and 37 ℃. In addition, the mutants of *Afe*KDO8PSs analysed in this chapter had previously been created by site-directed mutagenesis. Some of the mutant enzymes had also already been characterised, the results of which are included to aid discussion and are clearly indicated. Some experiments performed with *Afe*KDO8PS in this chapter were also performed in parallel on *Nme*KDO8PS. Like for *Afe*KDO8PS, the gene for *Nme*KDO8PS had been previously cloned into an expression vector, a purification protocol developed and the enzyme characterised.[39,81]

In this chapter:

- The previous characterisation of wild-type *Afe*KDO8PS is extended;
- Whether a metal-independent variant can be generated for the mesophilic *Afe*KDO8PS is investigated;
- Mutants of the SDGP motif in *Afe*KDO8PS are examined to probe the role of these residues and of the metal ion in catalysis.

The results, in conjunction with those on other KDO8PSs, point to a central role of the metal ion in positioning the β8α8 loop (bearing the SDGP motif) for catalytically productive A5P binding.

3.2 Residues chosen for mutation

Alignment of the *Afe*KDO8PS amino acid sequence with structurally characterised enzymes allowed identification of the key active-site residues (Figure 3.3). The C21N mutant of *Afe*KDO8PS was generated, as the presence of an Asn rather than a metal-binding Cys is the key difference between metal-independent and metal-dependent forms of KDO8PS. In addition, the role of the metal-binding Asp243 was examined by generation of D243A and D243E mutant enzymes. Mutant enzyme P245A *Afe*KDO8PS was also generated as this Pro residue is poorly conserved in the metal-dependent KDO8PSs.

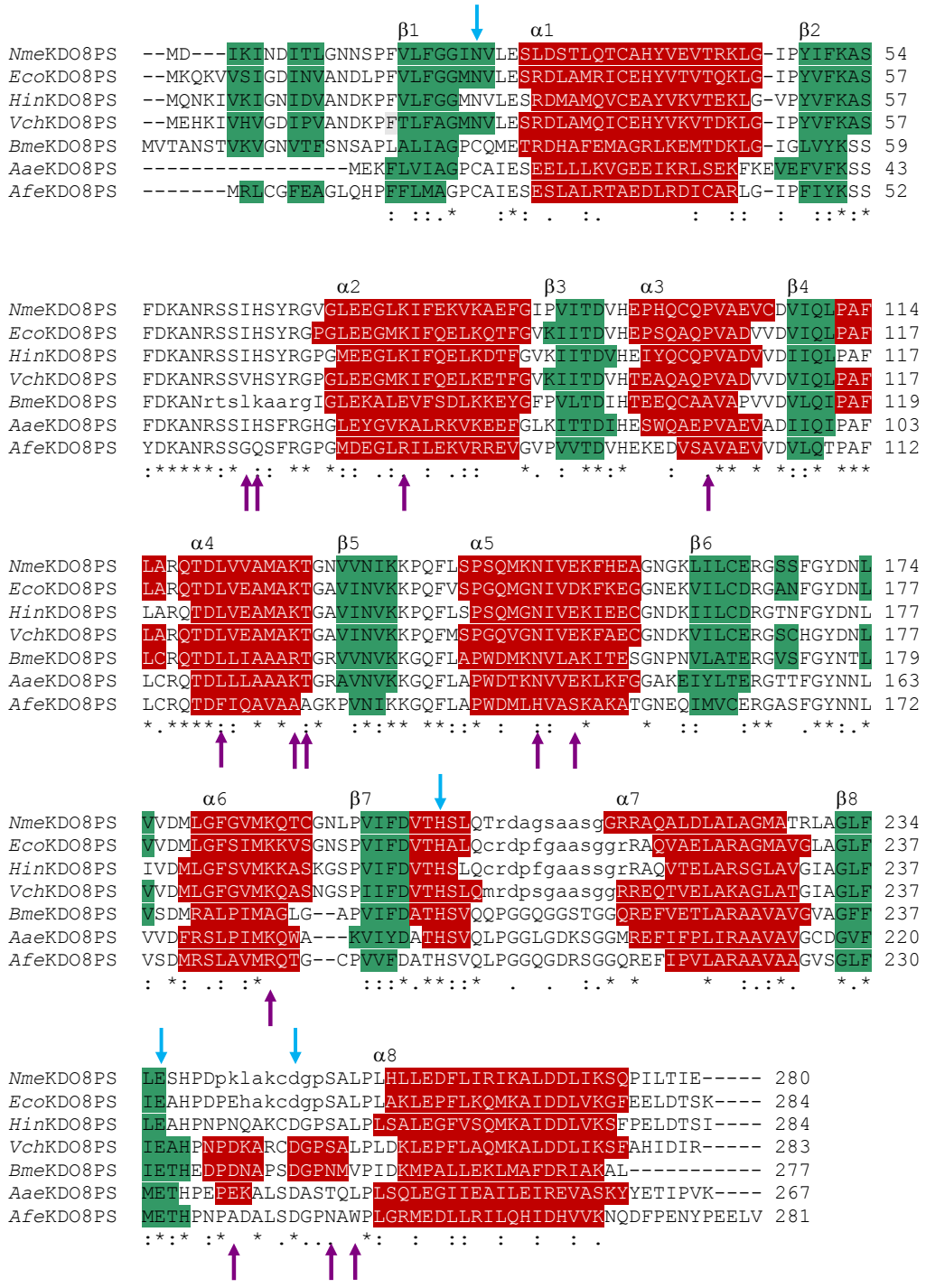


Figure 3.3: Structure-based alignment of KDO8PSs. The sequence of *Afe*KDO8PS was structurally aligned with those of structurally characterised KDO8PSs by means of PsiPred and pGenTHREADER.[91,92] Metal-binding residues are indicated with blue arrows. Residues in *Afe*KDO8PS uncommon to KDO8PSs are indicated by purple arrows. Residues that were not observed in the X-ray structures are shown in lower case. The character of sequence regions are colour coded: α-helical red and β-sheet green. *Hin*: *H. influenzae*, *Vch*: *V. cholerae*, *Bme*: *B. melitensis*.

3.3 Expression, purification and protein characterisation

The open-reading frame for the gene corresponding to *Afe*KDO8PS was previously cloned into a pT7-7 expression vector by Yeoman.[84] Lysis conditions and a two-step purification protocol, consisting of anion-exchange chromatography (AEC) followed by hydrophobic-interaction chromatography (HIC), had been developed. These are the same purification steps used for *Nme*KDO8PS.[39] For both *Afe*KDO8PS and *Nme*KDO8PS, a third size-exclusion chromatography (SEC) purification step was added after HIC to further increase the purity of the protein samples and desalt the proteins. Mutant *Afe*KDO8PS proteins behaved similarly to wild type and were expressed and purified using the same methods. Molecular weights of mutant proteins determined by mass spectrometry are listed in Table B.1.

Yeoman had found that approximately half of the over-expressed *Afe*KDO8PS was insoluble after cell lysis. As an alternative to sonication, cell lysis by chemical detergent (BugBuster®) was trialled. The supernatant liquid after centrifugation of the BugBuster®-lysed cells tended to be very viscous and difficult to filter (0.2 µm), despite using the additive Benzonase® Nuclease (an engineered promiscuous endonuclease). The relative soluble fraction of *Afe*KDO8PS was not improved by this alternate cell lysis method, suggesting perhaps that lysis was not the cause of insolubility.

Protein over-expression by the auto-induction method of Studier[93] was trialled and used for the production of some *Afe*KDO8PS protein batches. This method produced a much larger cell pellet for the same size growth volume compared to induction using isopropyl β-D-1-thiogalactopyranoside (IPTG), but that is likely due to the extended overnight incubation time compared to the 4 h induction period of the latter. As a result, a larger amount of protein is produced, but the crude lysate (supernatant liquid after centrifugation post-lysis) was viscous and difficult to handle. The viscosity did not appear to be proportional to the increased cell pellet size, and may

be caused by some other property of the autoinduction method. As such, protein over-expression induced by IPTG was preferred and this was the most often used method.

3.4 Metal activation

The apo-enzyme of *Afe*KDO8PS (purified in the presence of 1 mM EDTA) has no detectable activity when 10 mM EDTA is included in the assay, but activity could be restored by the addition of various divalent metal ions (Figure 3.4). A variety of divalent metal ions were tested for their ability to activate apo-*Afe*KDO8PS. Mn^{2+}, Co^{2+} and Cd^{2+} were the most activating divalent metal ions for *Afe*KDO8PS in order of decreasing activation (Figure 3.4). The metal-dependent *Apy*KDO8PS[65] was shown to possess maximal activity in the presence of Mn^{2+} while also being strongly activated by Co^{2+} and Cd^{2+}. Cd^{2+} was the most activating metal ion for *Aae*KDO8PS[82] and

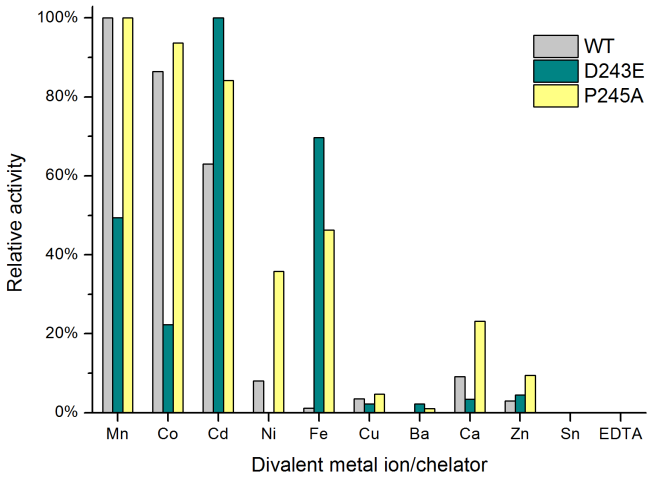


Figure 3.4: Metal-activation profiles of wild-type *Afe*KDO8PS, *Afe*D243E and *Afe*P245A. Relative activity has been normalised to the highest activating metal ion for each enzyme. Results for *Afe*D243E and *Afe*P245A are those of Yeoman.[84] Mg^{2+} was not an activating metal ion for wild-type *Afe*KDO8PS.

H. pylori KDO8PS (*Hpy*KDO8PS).[41,66]

As part of this thesis the metal-dependencies of wild-type and C21N *Afe*KDO8PS were assessed, while the three other mutant enzymes (D243A, D243E and P245A) were kinetically characterised by Yeoman.[84] C21N and D243A *Afe*KDO8PS demonstrated almost undetectable KDO8PS activity in the presence or the absence of metal ion or EDTA. On the other hand, the D243E and P245A mutants demonstrated metal-dependent activity. However, altered patterns of metal-dependency were observed for these latter two mutants in comparison to wild-type enzyme (Figure 3.4). The wild-type and P245A *Afe*KDO8PS were most active in the presence of Mn^{2+}, whereas for the D243E mutant the activity of the Cd^{2+}-activated enzyme was more than double the activity in the presence of Mn^{2+}. The P245A mutant showed less specificity for metal-ion type. In contrast to the wild-type enzyme, which displayed around 8 percent activity in the presence of Ni^{2+} or Fe^{2+}, the P245A mutant had significant activity (36 percent and 46 percent respectively) in the presence of Ni^{2+} and Fe^{2+}. Interestingly, Ni^{2+} was unable to activate the D243E mutant.

3.5 Kinetic properties

The values of the kinetic parameters for wild-type *Afe*KDO8PS were determined in the presence of each of the two most activating divalent metal ions, Mn^{2+} and Cd^{2+}. In the presence of Mn^{2+}, $K_m^{\text{PEP-Mn}}$ is $12 \pm 1\,\mu M$, $K_m^{\text{A5P-Mn}}$ is $22 \pm 2\,\mu M$ and k_{cat}^{Mn} is $4.8 \pm 0.1\,s^{-1}$. Minimal changes to these values were observed when Cd^{2+} replaced Mn^{2+} in the assay ($K_m^{\text{PEP-Cd}}$ is $13 \pm 1\,\mu M$, $K_m^{\text{A5P-Cd}}$ is $36 \pm 2\,\mu M$, and k_{cat}^{Cd} is $3.40 \pm 0.06\,s^{-1}$. These values are similar in magnitude to those observed for KDO8PS enzymes from *N. gonorrhoeae*,[44] *H. pylori*,[41,66] *N. meningitidis*,[81] *E. coli*[15] and *S. typhimurium*[46] (25 °C to 37 °C). The hyperthermophilic *Aae*KDO8PS and *Apy*KDO8PS show larger K_m^{PEP} and K_m^{A5P} values, but similar k_{cat} values (to *Afe*KDO8PS) at 90 °C and 60 °C respectively.[40,65] The very low level of activity of C21N *Afe*KDO8PS

Table 3.1: Kinetic parameters for Mn^{2+}-activated wild-type *Afe*KDO8PS and mutants. Those for D243E, D243A and P245A were determined by Yeoman.[84]

*Afe*KDO8PS	K_m^{A5P}	K_m^{PEP}	k_{cat}	k_{cat}/K_m^{A5P}	k_{cat}/K_m^{PEP}
Wild-type	21 ± 2	12.0 ± 0.7	4.80 ± 0.07	0.23 ± 0.03	0.40 ± 0.03
D243E	2630 ± 30	13 ± 2	4.8 ± 0.3	0.0018 ± 0.0001	0.37 ± 0.08
P245A	24 ± 5	5.2 ± 0.5	4.4 ± 0.2	0.18 ± 0.05	0.85 ± 0.10
D243A	–	–	0.005 ± 0.001	–	–
C21N-Mn^{2+}	–	–	0.004 ± 0.001	–	–
C21N-EDTA	–	–	0.004 ± 0.001	–	–

K_m values are in units of µM, k_{cat} in s^{-1}, and k_{cat}/K_m in $s^{-1}\,\mu M^{-1}$.

was unchanged by the addition of EDTA or by the addition of Mn^{2+}, Co^{2+} or Cd^{2+}.

The alternative aldose phosphate substrates 2-deoxyribose 5-phosphate (2dR5P) and R5P were not substrates for *Afe*C21N, regardless of metal-ion presence (or absence). The D243A *Afe*KDO8PS mutant was also barely active even with very high levels of substrates and high enzyme concentrations. In contrast, D243E and P245A *Afe*KDO8PS mutants were active (Table 3.1). Both these enzymes displayed k_{cat} values similar to that of wild-type enzyme. The most dramatic effect on the enzyme kinetics was the large elevation of K_m^{A5P} for the D243E mutant.

3.6 Structure in solution and stability

Sedimentation velocity experiments were performed for both wild-type and C21N *Afe*KDO8PS by analytical ultra-centrifugation (Figure 3.5). Data were fitted to a continuous size $[c(s)]$ distribution model, which showed a single peak for both proteins at a similar sedimentation coefficient value (Figure 3.5c). The peaks correspond to a molar mass of 126 kDa and 117 kDa for wild-type and C21N *Afe*KDO8PS respectively, which agrees with the calculated mass of the tetrameric enzyme (122 kDa). The same data were also collected at lower protein concentrations, which indicated similar sedimentation coefficients.

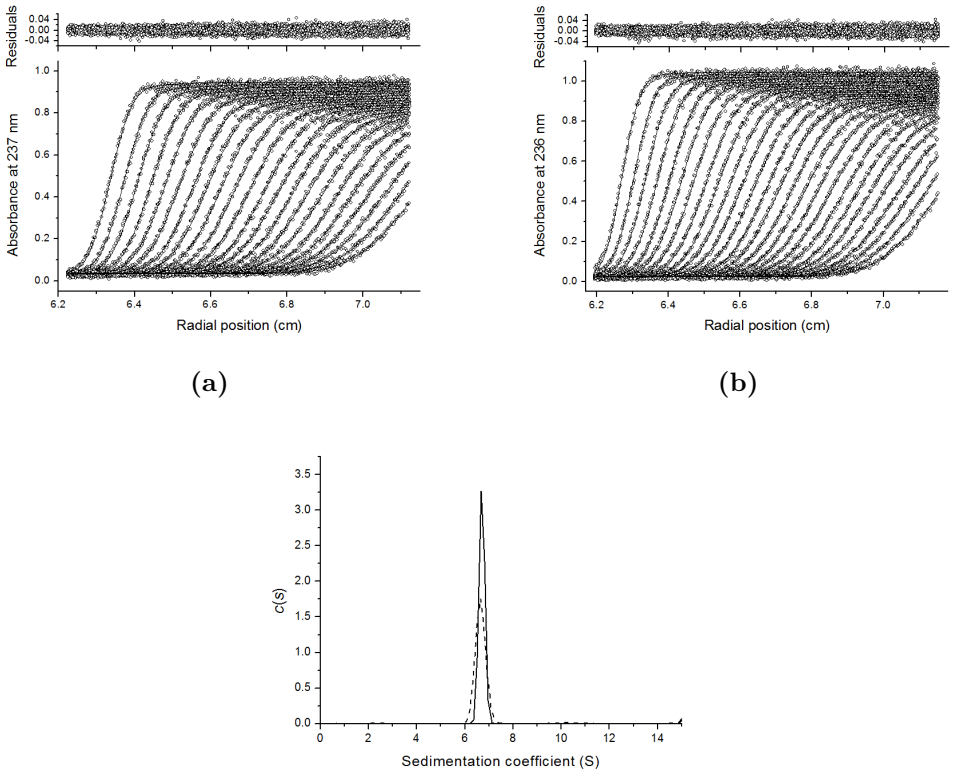


Figure 3.5: Sedimentation velocity analysis of wild-type and C21N *Afe*KDO8PS. The residuals (top pane) and absorbance as a function of radius (bottom pane) of (a) wild-type *Afe*KDO8PS and (b) C21N *Afe*KDO8PS. Every second scan is plotted superimposed with the non-linear least-squares fit (solid line) to a $[c(s)]$ model. (c) Overlay of the continuous size-distribution analyses for wild-type (solid line) and C21N (dashed line) *Afe*KDO8PS.

This determination of a tetrameric quaternary structure agrees with analysis on wild-type protein by analytical SEC.[84] All KDO8PSs studied to date have been tetramers in both crystal form and solution as suggested by analytical SEC.[39,40,45,46]

Differential scanning fluorimetry (DSF) was used to investigate the thermal stability of wild-type *Afe*KDO8PS and selected mutants. DSF was performed with combinations of three activating metal ions and the substrates PEP and A5P (Table 3.2). In the presence of 80 µM Mn^{2+} or Co^{2+} the T_m

of wild-type *Afe*KDO8PS was relatively unaffected, with values of 68 ℃ to 72 ℃ and 72 ℃ to 75 ℃ respectively. In the presence of 80 µM Cd^{2+} the T_m was raised to between 82 ℃ and 83 ℃. The presence of PEP and A5P had no significant effect on the measured T_m, except in the case of Cd^{2+} and PEP, and Co^{2+} and A5P. The thermal stability of the C21N *Afe*KDO8PS mutant was unchanged relative to wild type and not affected by the addition of 80 µM Cd^{2+}.

The thermal stability of the metal-independent *Nme*KDO8PS was also measured with a variety of additives. For this enzyme T_m values ≈10 ℃ below the lowest values determined for the metal-dependent *Afe*KDO8PS were recorded (Table 3.2). Intriguingly, whereas addition of Cd^{2+} enhanced the stability of the metal-dependent *Afe*KDO8PS and did not affect that of the metal-independent C21N *Afe*KDO8PS, in the presence of Cd^{2+} the T_m of *Nme*KDO8PS was lowered by 16 ℃. Cd^{2+} has been observed to strongly inhibit *Nme*KDO8PS at micromolar concentrations.[39] The explanation for this inhibition is unclear, as no bound metal ion was observed for Cd^{2+}-treated *Nme*KDO8PS crystals. However, it is interesting to note that on conversion of this enzyme into an obligate metal-dependent enzyme, the Cd^{2+} inhibition is lost, suggesting that inhibition may result from binding in a manner that disrupts the active site.[39]

The secondary structures of wild-type and C21N *Afe*KDO8PS and *Nme*KDO8PS were investigated using CD spectrophotometry. The spectra recorded for wild-type and C21N *Afe*KDO8PS were identical to one another and also to the spectrum for the structurally characterised *Nme*KDO8PS, indicating that *Afe*KDO8PS (and the C21N mutant) shares a similar secondary structure with *Nme*KDO8PS (Figure 3.6), and, given sequence similarity and enzyme activity profiles, a very similar $(\beta/\alpha)_8$ tertiary structure. CD was also used to measure the T_m for both wild-type *Afe*KDO8PS and C21N *Afe*KDO8PS by monitoring the decrease in ellipticity at 220 nm while the temperature was simultaneously increased. Similar to the results found by DSF, the T_m for C21N *Afe*KDO8PS was unchanged from the value calculated for wild-type of 73 ℃ (Figure 3.7).

Table 3.2: The effect of additives on T_m measured by DSF for wild-type *Afe*-KDO8PS and mutants, and wild-type *Nme*KDO8PS.

KDO8PS	Additives	T_m (°C)	Difference
*Afe*KDO8PS	No additive	68 ± 2	
	Mn^{2+}	68 ± 2	0
	Cd^{2+}	82 ± 1	14
	Co^{2+}	72 ± 2	4
	PEP + Mn^{2+}	72 ± 1	4
	PEP + Cd^{2+}	83 ± 1	15
	PEP + Co^{2+}	72 ± 1	4
	A5P + Mn^{2+}	69 ± 2	1
	A5P + Cd^{2+}	82 ± 2	14
	A5P + Co^{2+}	75 ± 1	7
	PEP	69 ± 1	1
	A5P	69 ± 1	1
*Nme*KDO8PS	No additive	58 ± 1	
	Mn^{2+}	59 ± 1	1
	Cd^{2+}	42 ± 1	−16
	Co^{2+}	57 ± 1	−1
	PEP	60 ± 1	2
	A5P	58 ± 1	0
*Afe*C21N	No additive	69 ± 1	
	Mn^{2+}	69 ± 1	0
	Cd^{2+}	69 ± 1	0
	Co^{2+}	69 ± 1	0
	PEP + Mn^{2+}	69 ± 1	0
	PEP + Cd^{2+}	70 ± 1	1
	PEP + Co^{2+}	69 ± 1	0
	PEP	70 ± 1	1
*Afe*P245A	No additive	71 ± 1	
	Mn^{2+}	71 ± 1	0
	Cd^{2+}	79 ± 2	8
	Co^{2+}	71 ± 1	0
	PEP + Mn^{2+}	71 ± 1	0
	PEP + Cd^{2+}	79 ± 1	8
	PEP + Co^{2+}	72 ± 1	1
	PEP	70 ± 1	−1
*Afe*D243E	No additive	68 ± 1	
	Mn^{2+}	70 ± 1	2
	Cd^{2+}	78 ± 1	10
	Co^{2+}	69 ± 1	1
	PEP + Mn^{2+}	70 ± 1	2
	PEP + Cd^{2+}	77 ± 1	9
	PEP + Co^{2+}	70 ± 1	2
	PEP	70 ± 1	2

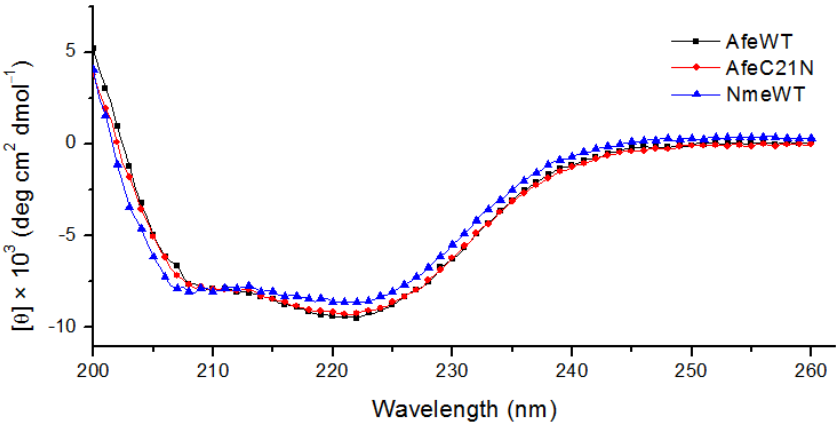


Figure 3.6: CD spectrophotometry of wild-type and C21N *Afe*KDO8PS (black squares and red circles respectively) and wild-type *Nme*KDO8PS (blue triangles). Every second datum point is plotted.

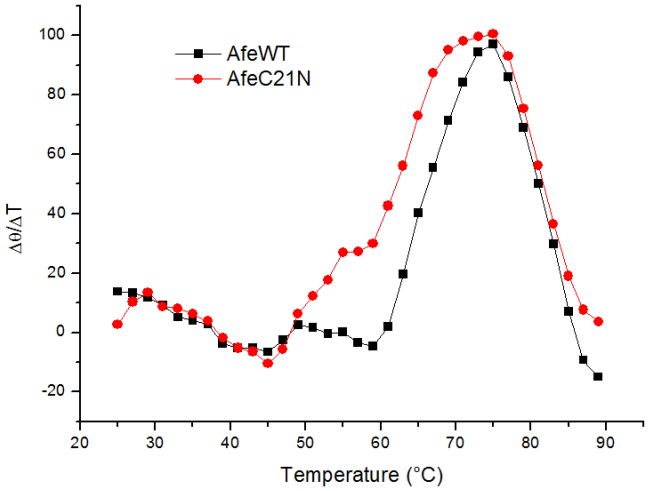
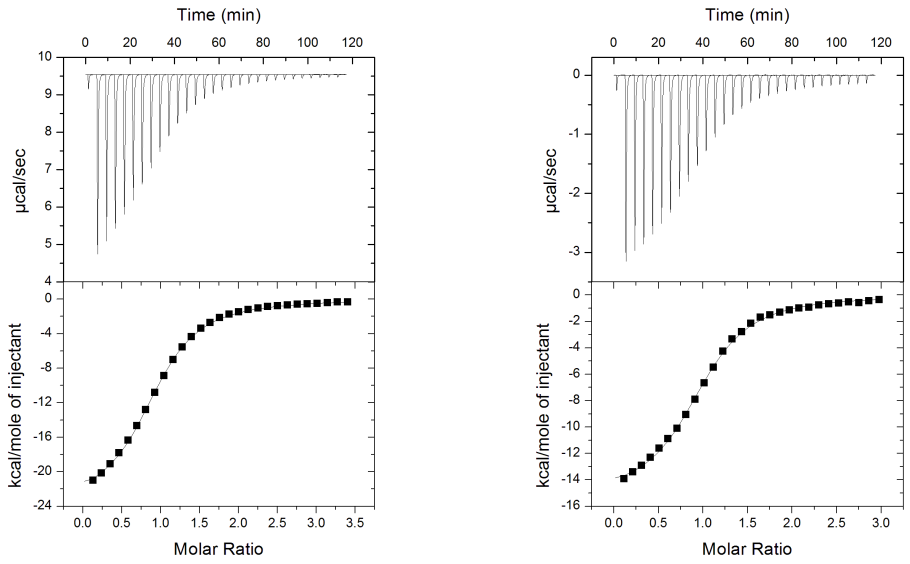


Figure 3.7: T_{m} determination of wild-type (black squares) and C21N (red circles) *Afe*KDO8PS by CD. T_{m} values for wild-type and C21N *Afe*KDO8PS were both calculated to be 73 °C.

3.7 Substrate specificity and binding affinities

Tight substrate specificity is a feature of KDO8PSs. The substrate specificity of *Afe*KDO8PS was examined by testing a range of phosphorylated aldoses as alternatives for A5P both kinetically and by ITC. ITC was used to determine the dissociation constants of both the metal-dependent *Afe*KDO8PS and metal-independent *Nme*KDO8PS for the substrates A5P and PEP, and the analogue R5P in the presence of PEP. PEP bound to *Afe*KDO8PS with a K_D of 5.0 ± 0.1 µM and a stoichiometry of 0.932 ± 0.004 *Afe*KDO8PS monomer per A5P molecule (Figure 3.8a). The corresponding values for *Nme*KDO8PS are a K_D of 4.8 ± 0.2 µM and a stoichiometry of 0.989 ± 0.006 (Figure 3.8b). Compared to PEP binding, both *Afe*KDO8PS and *Nme*KDO8PS yielded higher dissociation constants for A5P binding of 152 ± 2 µM and 29 ± 3 µM



(a) **(b)**

Figure 3.8: Interaction of PEP with (a) *Afe*KDO8PS and (b) *Nme*KDO8PS, quantified by ITC.

respectively (Figure 3.9a and 3.9b). Binding of the inactive A5P analogue R5P (*vide infra*) in the presence of PEP (150 µM) showed no heat change for *Afe*KDO8PS, suggesting R5P is not binding to the enzyme in the presence of PEP. In the case of *Nme*KDO8PS a K_D of 45 ± 3 µM was observed for R5P binding in the presence of PEP (Figure 3.9c).

No PEP consumption in standard kinetic assays was observed when D-glucose 6-phosphate or E4P was used in place of A5P in the assay mixture, even at significantly higher aldose and enzyme concentrations. 2dR5P was found to be a poor substrate for wild-type *Afe*KDO8PS, with K_m of 5.0 ± 0.8 mM, k_{cat} of 0.66 ± 0.06 s^{-1}, and k_{cat}/K_m of (1.4 ± 0.3) × 10^{-4} s^{-1} μM^{-1}. In contrast, 2dR5P was a considerably poorer substrate for the metal-independent *Eco*KDO8PS[80] and *Nme*KDO8PS enzymes[69] (55- and 20-fold decrease in k_{cat} respectively) than for *Afe*KDO8PS (seven-fold decrease). In the presence of R5P, loss of PEP was observed at a rate of 0.005 s^{-1}, 0.1 percent of that in the presence of A5P. R5P is not a substrate for the metal-independent *Nme*KDO8PS[69] and *Eco*KDO8PS,[80] nor the metal-dependent *Aae*KDO8PS.[27]

3.8 Crystallisation

Previous attempts to find crystallisation conditions for *Afe*KDO8PS had been made by using Crystal Screen I and II (Hampton Research).[84] Promising crystals grew in four conditions: crystals from two of the conditions did not diffract, and the crystals from the other two conditions diffracted poorly. The latter two conditions were further explored but crystal form and diffraction could not be improved.

In this thesis, *Afe*KDO8PS was screened in combinations of the absence of a metal ion, Mn^{2+}, Cd^{2+} and PEP, and at both room temperature and 8 °C. The screens and conditions initially trialled were:

- Structure screen I and II (Molecular Dimensions), without metal ion, with Mn^{2+} and with Cd^{2+}

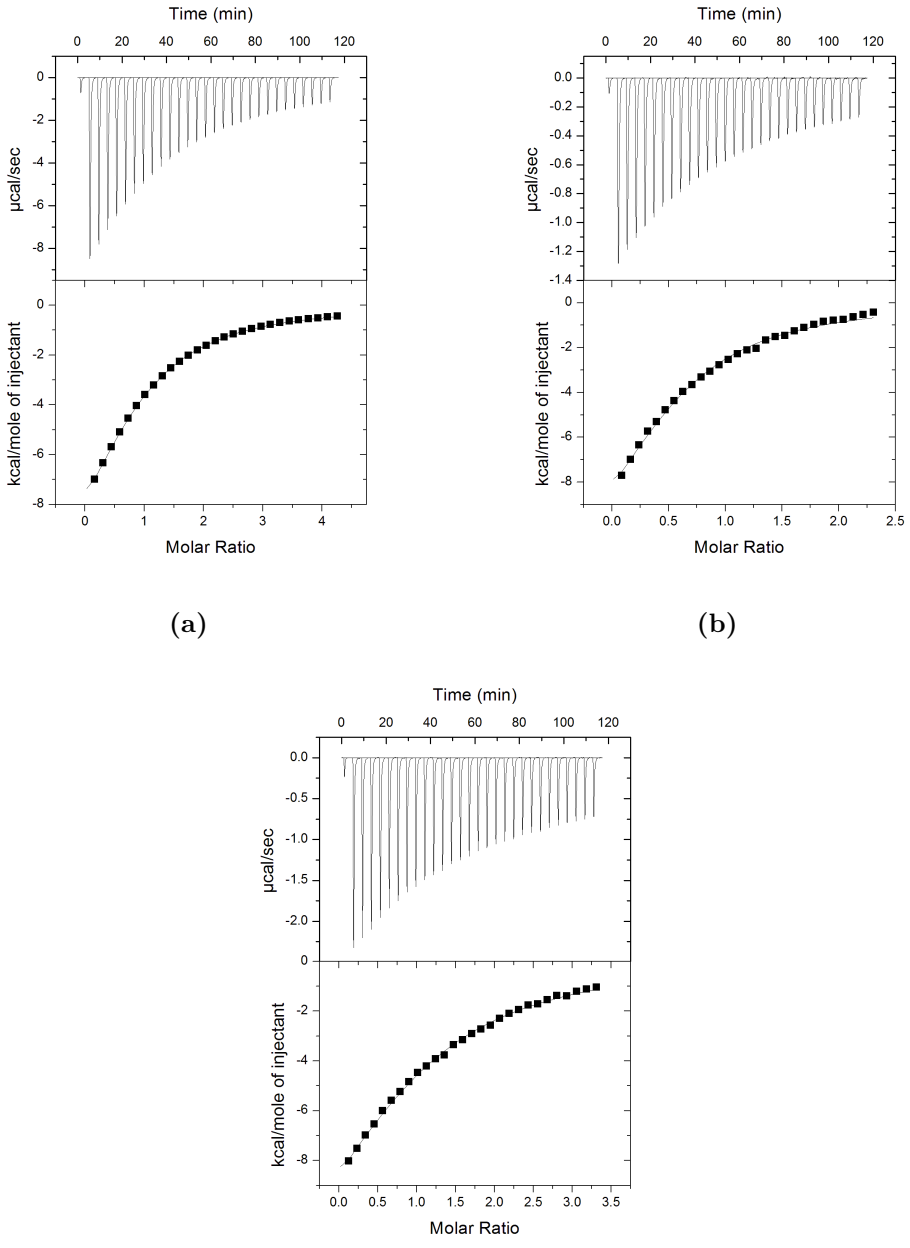


Figure 3.9: Interaction of A5P with (a) *Afe*KDO8PS and (b) *Nme*KDO8PS quantified by ITC. (c) Interaction of R5P with *Nme*KDO8PS in the presence of PEP.

- Collaborative Crystallisation Centre (C3) screen 2 (PEG-based) with Cd^{2+} and PEP
- C3 screen 3 (PEG-based) with Cd^{2+} and PEP
- C3 screen 4 (PEG-based) with Cd^{2+} and PEP
- C3, a random screen of PEG-based conditions with Cd^{2+} and PEP
- PACT, a PEG/ion/pH screen, at C3 and locally
- JCSG+ suite, a sparse matrix screen, at C3 and locally

Crystals of varying visual quality and size grew in many conditions (Figure 3.10). The top-hit crystals were tested for diffraction quality at either the Institute of Fundamental Sciences, Massey University or the MX1 and MX2 beamlines at the Australian Synchrotron. For all crystals, diffraction was to a poor resolution (≈8 Å) and highly mosaic, or there was no diffraction. Crystals from two different conditions diffracted to a resolution of ≈3.5 Å, and datasets were collected. Unfortunately it was not possible to integrate the data for both sets. Crystal conditions were further optimised for many of the original hit conditions by varying growth temperature, component concentrations of the crystallisation condition, metal ion and PEP. While crystals could be reproduced of similar morphology to those from the original screens, none of the optimised conditions resulted in better diffracting crystals.

Crystallisation condition screening of *Nme*KDO8PS

The C3 facilities were also used to screen *Nme*KDO8PS for new crystallisation conditions. The known condition (100 mM sodium acetate pH 4.6, 0.6 M to 4.0 M NaCl) produces crystals readily for *Nme*KDO8PS over a large range of protein concentrations (2 mg mL^{-1} to 30 mg mL^{-1}). However, the low pH of the condition is believed to hinder ligand binding due to protonation of carboxylate, histidyl and phosphate groups.[39] New conditions, ideally at a pH more amenable to ligand binding were sought that would allow co-crystallisation with or soaking in of ligands. *Nme*KDO8PS was screened with

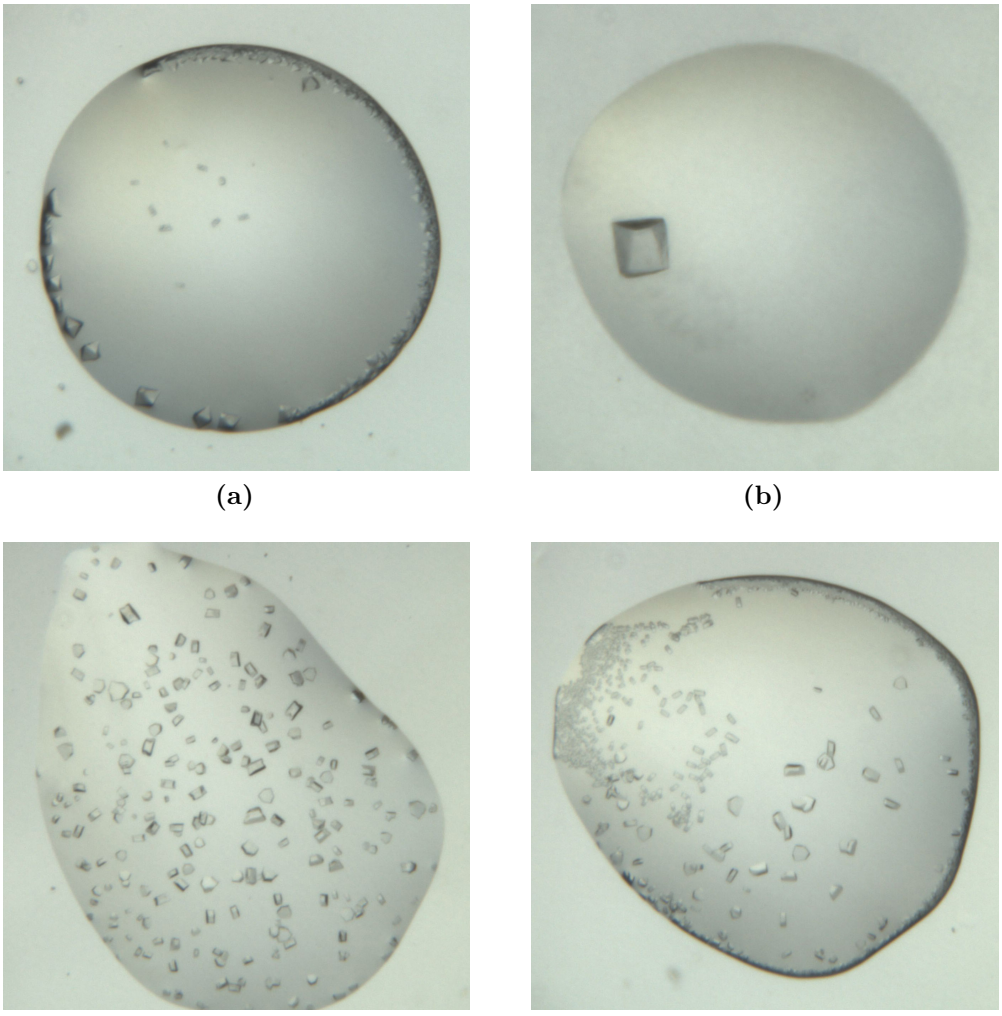


Figure 3.10: Photographs of *Afe*KDO8PS crystals from crystallisation trials at the C3. (a) PACT screen, condition G7. (b) C3-2 screen, condition E8. (c) C3-2 screen, condition A7 (d) C3-2 screen, condition B11.

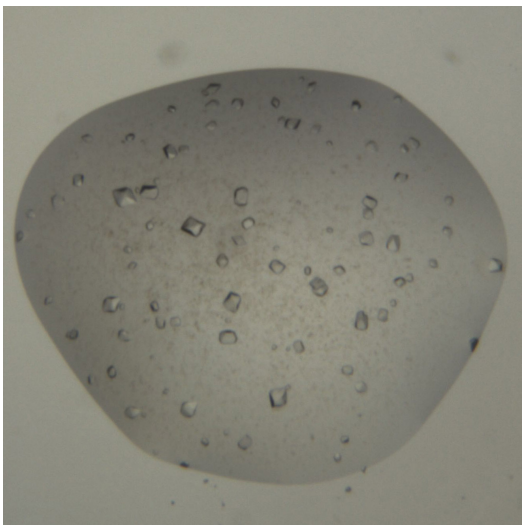


Figure 3.11: Photograph of drop containing crystals of *Nme*KDO8PS. The drop corresponds to condition E10 from the C3 screen C3-3.

2 mM PEP, at both 8 °C and 20 °C in C3 screens 2, 3 and 4, and at 20 °C in the PACT screen. The only condition in which crystals grew was in C3 screen 3, in a condition composed of 100 mM sodium acetate pH 5.0, 1 M NaCl and 10 % (w/v) PEG 4000 (Figure 3.11). This condition is very similar to the condition already known and the crystals that grew were of a similar morphology. No further attempts were made to find new crystallisation conditions for this protein.

3.9 Discussion

Active wild-type *Afe*KDO8PS was heterologously expressed and purified from *E. coli*. In general, this enzyme exhibits similar properties to other characterised KDO8PSs, having comparable kinetic parameters, a narrow substrate specificity, a similar fold, and similar tetrameric quaternary structure. As predicted from its sequence, which demonstrates the presence of a metal-binding Cys, *Afe*KDO8PS displays activation by divalent metals. However, the work in this thesis and that by Yeoman[84] highlights significant differences shown by this KDO8PS compared to those from other organisms.

Mechanistic studies of KDO8PS have focused on the order of attack of the water on C2 of PEP and attack of PEP on C1 of A5P, and on the mechanisms that activate the water for nucleophilic attack and the carbonyl carbon for electrophilic attack (Figure 3.12).[19,73] In this regard there are a number of possibilities for the role the metal ion plays in metal-dependent KDO8PSs. The metal ion may play a role in directly activating the carbonyl functionality of A5P (Figure 3.12a), or may play a role in the activation of the nucleophilic water (Figure 3.12b). In addition, the metal ion may play a structural role whereby the coordination to the Asp ligand (Asp243 in *Afe*KDO8PS) may enable this residue to then participate directly in the reaction or in the binding of substrates.[73]

Although *Afe*KDO8PS was strongly activated by Cd^{2+} the enzyme is unlikely to encounter this metal ion in vivo. The second most abundant intracellular divalent metal ion Mg^{2+} was not activating for *Afe*KDO8PS. Metal ion content analyses of other KDO8PS enzymes has suggested that in vivo, Fe^{2+} and Zn^{2+} are bound in equal measure to KDO8PS.[65,82] These two metal ions were minimally activating for *Afe*KDO8PS.

3.9.1 D243E, D243A and P245A *Afe*KDO8PS mutants

A notable feature of the active D243E and P245A *Afe*KDO8PS mutants is the altered patterns of metal activation in comparison to wild-type enzyme (Figure 3.4). Substitution of Pro245 for Ala may provide conformational freedom to the main chain allowing a greater range of metal ions to be accommodated, while still presenting the metal-ligand Asp243 in the conformation required by the A5P substrate.

Apart from the difference in metal-activation patterns, the mutation P245A caused little alteration to the kinetic parameters with respect to A5P and PEP. This Pro is two residues downstream of the Asp metal-binding ligand, which is found on the β8α8 loop in the structures of other KDO-

8PSs (Figure 3.3). In metal-independent KDO8PSs this Pro is invariant. Mutation of this residue to Ala in the metal-independent *Nme*KDO8PS also had no significant impact on the activity of this enzyme.[39] It seems likely that the invariance of this residue reflects the relatively recent evolution of the metal-independent enzymes from a metal-dependent KDO8PS ancestor.

In contrast, there was a substantial change of $K_{\mathrm{m}}^{\mathrm{A5P}}$ for the D243E mutant, with this parameter being over two orders of magnitude greater in the mutant enzyme when compared to the wild-type enzyme. The negligible activity of the D243A mutant is consistent with a critical role for this residue. In addition to metal binding, it has been suggested that this Asp may act to remove a proton from a metal-bound water in order to initiate the reaction with C2 of PEP.[73] In structures of *Aae*KDO8PS in which A5P is bound, this Asp has been shown to play a dual role in both metal-binding and in hydrogen bonding the C3 or C4 hydroxyl group of A5P (Figure 3.12 and 1.15). This role in A5P binding is consistent with the substantially elevated $K_{\mathrm{m}}^{\mathrm{A5P}}$ observed for the D243E *Afe*KDO8PS mutant. The dual role of Asp243 in metal and A5P binding may be sufficient to account for the lack of activity of the D243A mutant.

Altering the positioning of the carboxylate functionality, by substitution of Asp243 to Glu, also has a significant effect on the patterns of metal activation observed for this enzyme. In particular, Fe^{2+}, along with Cd^{2+}, is strongly activating for the D243E mutant. These two metal ions have the greatest preference for six-coordination, which might be accomplished by bidentate coordination of the conformationally more adjustable Glu residue compared to the Asp residue, while still preserving the key role of this residue in positioning the A5P substrate. It is interesting that Zn^{2+}, and also Ni^{2+} and Cu^{2+}, which have the strongest preference for four- or five-coordination are at best very weakly activating for this mutant, as well as for the P245A mutant and wild-type enzyme in the case of Cu^{2+} and Zn^{2+}. While it is tempting to infer from these data a Lewis-acid role for the metal ion in activating either a nucleophile for substrate-attack or direct activation of the aldehydic substrate (Figure 3.12a), the role of the metal ion, in conjunction

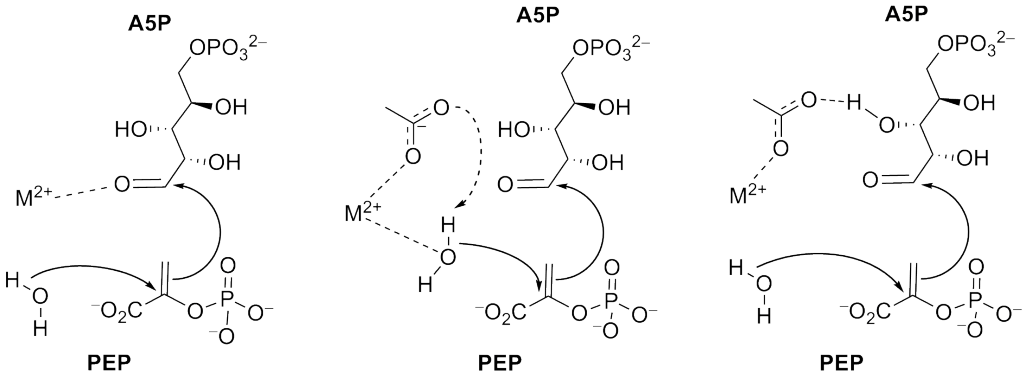


Figure 3.12: Possible roles for the metal ion in the KDO8PS reaction. (a) Metal activation by coordination to the A5P carbonyl. (b) Metal activation of nucleophilic water, with a possible role for the metal-binding Asp in water deprotonation. (c) Structural role in orientation of the metal-binding Asp to assist positioning of the β8α8 loop and substrate binding.

with the Asp ligand, may simply be to position correctly the A5P substrate for coupling with PEP,[29] as discussed below. This latter role is the most consistent with both forms of KDO8PS sharing a common mechanism and the ability in many cases for interconversion of metal dependency by single substitution mutagenesis.

3.9.2 C21N *Afe*KDO8PS mutant: role of metal ion in anchoring the β8α8 loop for catalysis

The almost undetectable activity of C21N *Afe*KDO8PS was surprising; therefore the correct folding of this mutant was assessed by CD spectrophotometry. In addition, the melting temperature determined for the C21N mutant was the same as the value obtained for wild-type enzyme. These findings indicate that the lack of catalytic activity observed for this mutant is not due to significant disruption of tertiary or quaternary structure. Importantly, there was also no measurable enthalpy of binding by ITC when C21N *Afe*KDO8PS was titrated with PEP, or A5P in the presence of PEP, suggesting the mutation

has severely compromised the ability of the enzyme to bind both substrates, resulting in the lack of activity.

The lack of activity of the C21N *Afe*KDO8PS is intriguing as the equivalent mutation produced somewhat active (and metal-independent) enzymes for the very closely related metal-dependent *Aae*KDO8PS and *Apy*-KDO8PS.[28,73] The ease with which the switch to metal-independency can be achieved is important for understanding the relationships between metal-dependent and metal-independent forms of KDO8PS. As well as possibly playing a direct role in the reaction catalysed by KDO8PS (Figure 3.12), coordination of a metal ion also helps define the position of the β8α8 loop (Figure 3.2). This loop bears the metal-binding Asp and Glu, and as noted above the Asp plays a dual role in metal coordination and A5P binding. The Asp, or at least its carboxylate functionality, is critical for catalysis for both metal-dependent and metal-independent KDO8PSs. In the structures of the metal-independent KDO8PSs from *E. coli* and *N. meningitidis* the β8α8 loop is generally poorly defined. In contrast this loop appears to be more precisely positioned for *Aae*KDO8PS, although this sharper definition may be an artefact of enhanced rigidity of a hyperthermophilic enzyme compared to mesophilic ones at the similar temperatures employed for crystallisation.

Furthermore, the characterised metal-dependent KDO8PSs for which the switch to metal-independency has been achieved by a single substitution of the metal-binding Cys-to-Asn are from hyperthermophiles. Hence, at the temperatures (≈60 °C) at which activity assays are performed, these KDO-8PSs are intrinsically more rigid than at their optimum growth temperatures of close to 100 °C, and, moreover, they are more rigid than *Afe*KDO8PS, which is mesophilic in its temperature preferences. Thus, in contrast to *Afe*KDO8PS, the metal-ion coordination may not be structurally essential for positioning the β8α8 loop in the KDO8PSs from hyperthermophiles, at least at assay temperatures. This is corroborated in the metal-free adducts of *Aae*KDO8PS where this loop remains observable, although more mobile, in electron-density maps,[73] and in the complete loss of activity in the D247A mutant of *Nme*KDO8PS (equivalent residue is the metal-binding Asp243 in

*Afe*KDO8PS and Asp233 in *Aae*- and *Apy*KDO8PS).[39] It is conceivable that if activity assays on the Cys-to-Asn mutants of *Aae*KDO8PS and *Apy*KDO8PS were performed at 100 °C, where flexibility is likely to be comparable to that of *Afe*KDO8PS at 37 °C, there would be only a very low level of activity, comparable to that seen for the C21N mutant of *Afe*KDO8PS.

For *Afe*KDO8PS, loss of ability to bind metal in the C21N mutant may mean that the β8α8 loop, which provides two of the metal-binding ligands, cannot be locked down, and/or that binding A5P is not sufficient to tie both A5P and this loop into a productive conformation for reaction of A5P with PEP. Flexibility of the β8α8 loop may also disrupt PEP binding. The absolutely conserved Glu on this loop (which acts as a bidentate metal-binding ligand in the metal-dependent KDO8PSs) is linked to the PEP carboxylate via a conserved Lys (Figure 3.13). Hence, if this loop is not held correctly in position, PEP binding would be disrupted, as observed by ITC with the C21N mutant.

The importance of the metal ion for activity is further emphasised by the extraordinary 14 °C increase in denaturation temperature in the presence of

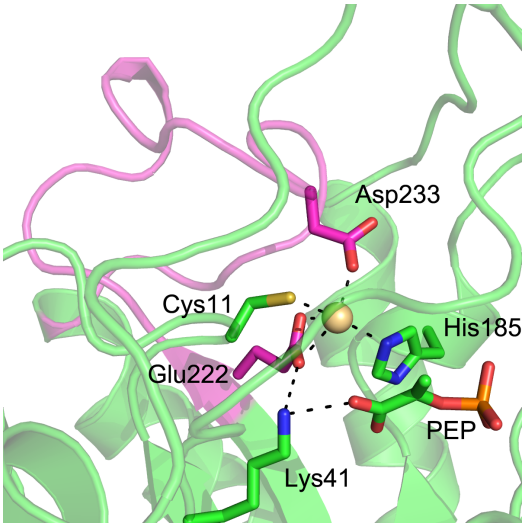


Figure 3.13: The structure of *Aae*KDO8PS (PDB code 1FWW). The carbon atoms of the β8α8 loop are coloured magenta, and the metal ion (Cd^{2+}) is coloured wheat.

Cd^{2+} for wild-type *Afe*KDO8PS, contrasted to no increase in stability in the presence of the slightly more activating Mn^{2+} ion. On the other hand, Cd^{2+} has a pronounced inhibitory and destabilising effect for the metal-independent *Nme*KDO8PS.[39]

Two of the highlighted differences between the two forms of KDO8PS remain unexplored for *Afe*KDO8PS. The first is the identity of the residue immediately before the mutated metal-ligand Cys21. As discussed earlier for *Eco*KDO8PS, in metal-independent KDO8PSs the two residues exist as a Met/Ile-Asn pair and in metal-dependent (and *Afe*KDO8PS), as Pro-Cys. In the conversion of *Eco*KDO8PS to become metal-dependent, the combination of M25P with N26C led to increased activity relative to the N26C mutation by itself. It may therefore be informative to assess the effects of combining the mutation P20M with C21N in *Afe*KDO8PS. If the Asn residue in metal-independent KDO8PSs fulfils the same role as the metal ion in coordinating the arrangement of the active site, then the identity of the preceding residue may be important for precisely controlling its position. The second unexplored difference is Ser242, which while required as a Ser in metal-dependent KDO8PSs to avoid formation of a disulfide bond with Cys21, in metal-independent KDO8PSs is conserved as a Cys, and has been suggested as one way in which this form compensates for the absence of metal ion. Mutation of both Pro20 and Ser242 could potentially be tried in future studies to achieve metal-independent activity for *Afe*KDO8PS.

3.10 Conclusions

*Afe*KDO8PS was expressed and purified and the existing initial characterisation expanded. This enzyme is metal-dependent, tetrameric, and displays the tight substrate specificity reported for other KDO8PSs. Several mutants of the C-terminal β8α8 loop of *Afe*KDO8PS were analysed, focusing on the mechanistic role of the metal-binding Asp in this loop. These results highlight a critical role for this loop in binding and positioning the A5P and PEP

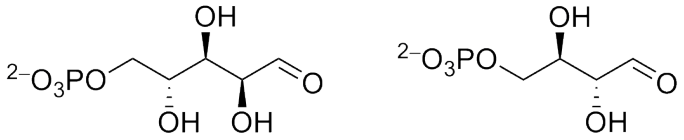
substrates. Specifically, in the absence of metal ion, as engineered by either the C21N or D243A mutations or by deprivation of metal ion in wild-type enzyme, there is minimal activity. Moreover, for the C21N mutant, ITC reveals an absence of both A5P and PEP binding. Mispositioning of the metal-binding Asp243 carboxylate functionality by mutation to Glu alters metal-ion preference, leaves k_{cat} unaffected for the Mn^{2+} form and greatly increases $K_{\mathrm{m}}^{\mathrm{A5P}}$, further reinforcing the crucial interdependence of metal ion and this loop. In contrast to the well-characterised *Aae*KDO8PS there appears to be no simple Cys/Asn switch associated with metal-dependency for *Afe*KDO8PS.

Chapter 4

The role of the KANRS motif for substrate selection and catalysis

4.1 Introduction

Many of the key details of the reactions catalysed by both KDO8PS and DAH7PS have been elucidated. It is known that the *si* face of PEP attacks the *re* face of the A5P or E4P aldehydic cosubstrate.[16] The reaction of PEP, A5P/E4P and water forms a linear tetrahedral intermediate, from which the phosphate ester of the PEP parent is subsequently eliminated through cleavage of the C-O bond.[29,94] However, what is less clear is how specific substrate selection is controlled by these enzymes. DAH7PSs have been shown to be relatively promiscuous and utilise, with variable efficiency, A5P (Figure 4.1a) and other five-carbon aldehyde substrates, as well as E4P (Figure 4.1b) and its variants, such as 2-deoxyerythrose 4-phosphate (2dE4P).[69,83] On the other hand, KDO8PSs exhibit tight specificity: the enzyme is not just specific for carbon length but also does not tolerate diastereoisomeric substrates in which the configurations of the hydroxyl groups at positions 2 and/or 3 are



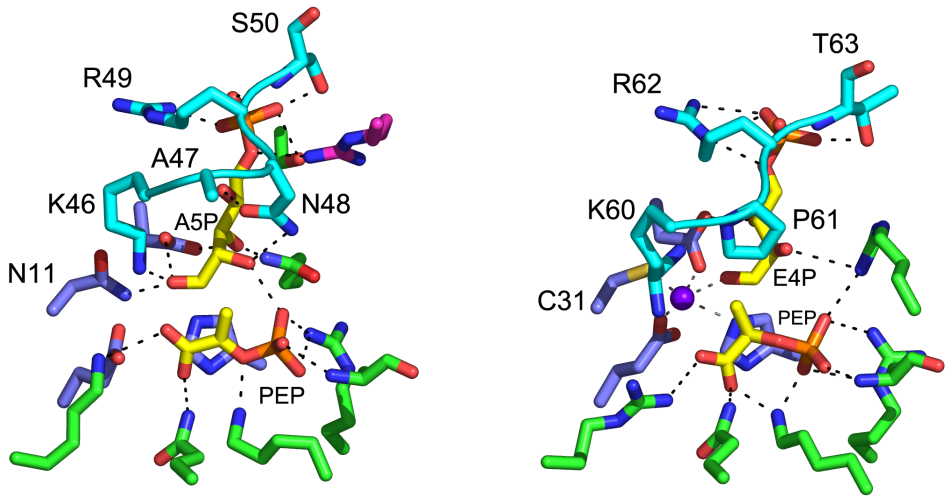
(a) A5P **(b)** E4P

Figure 4.1

inverted.[20,81]

Many of the residues that directly contact substrates are invariant for all KDO8PSs and DAH7PSs. Of particular note to substrate selection is an absolutely conserved motif of five amino acids in KDO8PSs that has been shown to interact directly with A5P. This LysAlaAsnArgSer (KANRS) motif is present on the β2α2 loop of the $(\beta/\alpha)_8$ barrel and is an integral part of the enzyme active site. In structures of *Aae*KDO8PS in which both PEP and A5P are bound, this motif is responsible for the key interactions of the enzyme with A5P (Figure 4.2).[24,72] The Arg and Ser residues coordinate to the phosphate moiety of A5P and form the phosphate binding site for the aldose substrate, along with an Arg contributed by the adjacent subunit, and a Ser from the β7α7 loop. The Asn residue of this KANRS motif forms hydrogen bonds to the C2 and C4 hydroxyl groups of A5P, providing contacts that may help to orient the reactive carbonyl of A5P in the active site. The Lys is positioned close to the A5P carbonyl and may play a role in its activation.[24] In DAH-7PSs, an absolutely conserved LysProArgSer/Thr (KPRS) motif occupies the same position as the KANRS motif of the KDO8PSs and is responsible for the interactions between the enzyme and substrate E4P. It is likely that this obvious conserved difference plays an important role in determining the differential substrate selection of the two enzymes.

In this chapter, the importance of the KANRS motif for catalysis and substrate selection by the metal-dependent *Afe*KDO8PS and metal-independent *Nme*KDO8PS has been examined. An essential catalytic role for the Lys of this motif has been disclosed and the importance of the Asn in substrate selection is illustrated.



(a) **(b)**

Figure 4.2: Binding of A5P and E4P to the active site of KDO8PS and DAH7PS. (a) *Aae*KDO8PS C11N/S235P/Q237A mutant (PDB code 2NX3). (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). The carbon atoms of the KANRS and KPRT motifs are coloured cyan, those of the metal-binding site residues blue, those of the PEP-binding residues green, those of residues from an adjacent subunit magenta, and those of PEP and A5P yellow. The manganese ion in panel (b) is coloured purple.

4.2 Preparation of A5P binding site mutants

The Lys residue of the KANRS motif is conserved in both the KDO8PSs and the related DAH7PSs. It has been proposed that the amino moiety of this residue plays a role in the protonation of the A5P carbonyl functionality in catalysis for KDO8PSs.[24] Mutant proteins were generated of both *Afe*-KDO8PS and *Nme*KDO8PS from which this functionality had been removed (K55A and K57A in *Afe*KDO8PS and *Nme*KDO8PS, respectively). The adjacent Ala and Asn residues represent the conserved difference between KDO8PSs and DAH7PSs, and these residues are likely to play a role in determining substrate selection. Mutant proteins were produced in which the amide functionality of the Asn was removed (N57A and N59A) and or altered

to a carboxylate moiety (N57D and N59D) to ascertain whether an isosteric hydrogen bond acceptor would suffice in this position. In addition, deletion mutant proteins were created by removing this Asn residue (KARS proteins) or by substituting both Ala and Asn with Pro, thereby creating the conserved KPRS motif of the DAH7PSs (KPRS proteins). The constructs for six of the mutant proteins (*Afe*K55A, *Nme*K57A, *Afe*N57A, *Afe*N57D, *Nme*N59A and *Nme*N59D) already existed and had been created by site-directed mutagenesis by Dr Fiona Cochrane (*Nme*KDO8PS) and Jeffrey Yeoman (*Afe*KDO8PS[84]).

Mutant proteins were purified according to procedures developed for the respective wild-type proteins.[39,84] The structural integrity for the mutants of *Afe*KDO8PS was assessed by CD spectrophotometry (Figure 4.3).

The melting temperatures of *Afe*KDO8PS and *Nme*KDO8PS mutant proteins were determined by DSF and were found to be similar to those measured for the respective wild-type proteins (Table 4.1). *Nme*KARS and *Nme*KPRS mutants both have overall lower melting temperatures than

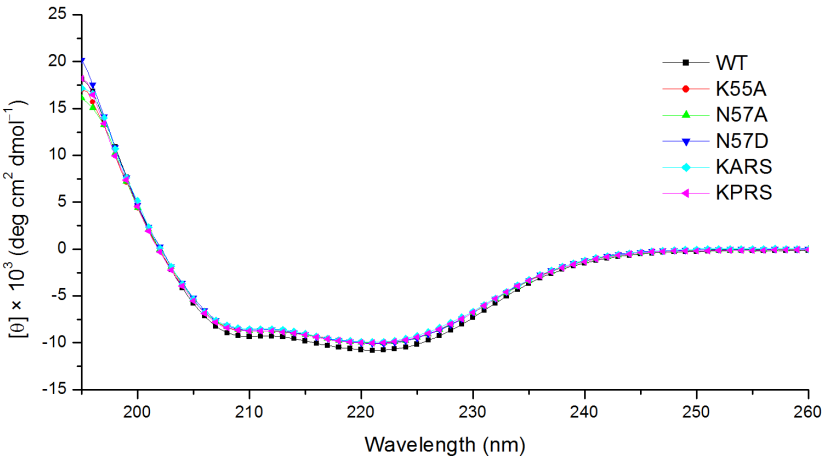


Figure 4.3: CD spectrophotometry of *Afe*KDO8PS wild type and mutants. Wild type: black squares; *Afe*K55A: red circles; *Afe*N57A: green triangles; *Afe*N57D: blue triangles; *Afe*KARS: cyan diamonds; *Afe*KPRS: pink triangles. Every second datum point is plotted.

Table 4.1: The effect of additives on T_m for wild-type and KANRS mutants of *Nme*KDO8PS and *Afe*KDO8PS.

		*Nme*KDO8PS		*Afe*KDO8PS	
KDO8PS	Additives	T_m (°C)	Diff.[a]	T_m (°C)	Diff.[a]
Wild-type	No additive	58 ± 1		68 ± 2	
	Mn^{2+}	59 ± 1	1	68 ± 2	0
	Cd^{2+}	42 ± 1	−16	82 ± 1	14
	Co^{2+}	57 ± 1	−1	72 ± 2	4
K57A/K55A	No additive	66 ± 0		68 ± 1	
	Mn^{2+}	66 ± 0	0	66 ± 1	−2
	Cd^{2+}	62 ± 1	−4	81 ± 1	13
	Co^{2+}	66 ± 1	0	68 ± 1	0
N59A/N57A	No additive	61 ± 1		70 ± 1	
	Mn^{2+}	60 ± 1	−1	69 ± 2	−1
	Cd^{2+}	47 ± 1	−14	80 ± 1	10
	Co^{2+}	61 ± 0	0	68 ± 2	−2
N59D/N57D	No additive	65 ± 1		68 ± 1	
	Mn^{2+}	65 ± 1	0	67 ± 1	−1
	Cd^{2+}	55 ± 1	−10	77 ± 1	9
	Co^{2+}	65 ± 1	0	68 ± 1	0
KARS	No additive	51 ± 3		70 ± 1	
	Mn^{2+}	54 ± 2	3	72 ± 1	2
	Cd^{2+}	NCT[c]	–	78 ± 2	8
	Co^{2+}	55 ± 1	4	70 ± 1	0
KPRS	No additive	48 ± 1		67 ± 1	
	Mn^{2+}	50 ± 1	2	68 ± 2	1
	Cd^{2+}	NCT	–	76 ± 1	9
	Co^{2+}	52 ± 2	4	69 ± 1	2

[a] Differences are with respect to no additive conditions.
[c] No clear transition

wild-type *Nme*KDO8PS, indicating that these mutations affect the inherent stability of the enzyme.

4.3 Mutations in the A5P binding site alter the kinetic profile of KDO8PS

The catalytic activities of all mutants were assessed (Table 4.2). Mutation of the KANRS motif Lys was completely deleterious to enzyme function, with both *Nme*K57A and *Afe*K55A having no perceptible activity. Compromised catalytic activity was observed for the mutants in which the Asn residue was altered to Ala. In accordance with the importance of this residue for A5P binding, $K_{\mathrm{m}}^{\mathrm{A5P}}$ was significantly increased, whereas the K_{m} values with respect to PEP were only slightly increased. However, the substitution of the amide moiety of this Asn by the carboxylate functionality (*Nme*N59D and *Afe*N57D) did not produce functional enzymes.

Contraction of the KANRS motif to KARS and its alteration to the DAH7PS-like KPRS motif also rendered the enzymes inactive. The only exception to this was the *Nme*KARS mutant; however, its maximal rate of reaction was more than 400 times slower than that observed for the wild-type enzyme, and the K_{m} values determined for this enzyme, particularly that for the use of A5P, were significantly higher than those observed for the wild-type enzyme.

Wild-type enzymes and mutants were also tested for their ability to accept alternative aldose phosphates as substrates to A5P. None of the enzymes exhibited any activity with E4P, the natural substrate for DAH7PS. R5P was similarly not accepted as an alternative. However, both wild-type enzymes and the compromised Asn mutants (*Nme*N59A and *Afe*N57A) were both able to accept 2dR5P as an alternative substrate (Table 4.3). Intriguingly, analysis of the specificity constants reveals that 2dR5P is a significantly better substrate than A5P for these active mutants than it is for wild-type enzymes.

Table 4.2: Kinetic parameters for wild-type and KANRS mutants of *Nme*KDO8PS and *Afe*KDO8PS.

KDO8PS	K_m^{PEP} (µM)	K_m^{A5P} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{A5P} ($s^{-1}\,mM^{-1}$)
*Nme*WT	2.5 ± 0.2	12.0 ± 0.5	8.0 ± 0.1	660 ± 40
*Nme*K57A	–	–	–	
*Nme*N59A	7.9 ± 0.7	503 ± 28	0.149 ± 0.002	0.29 ± 0.02
*Nme*N59D	–	–	–	
*Nme*KARS	43 ± 11	1100 ± 200	0.0178 ± 0.0007	0.016 ± 0.004
*Nme*KPRS	–	–	–	
*Afe*WT	12 ± 1	22 ± 2	4.8 ± 0.1	220 ± 20
*Afe*K55A	–	–	–	
*Afe*N57A	2.9 ± 0.6	81 ± 4	0.30 ± 0.01	3.7 ± 0.3
*Afe*N57D	–	–	–	
*Afe*KARS	–	–	–	
*Afe*KPRS	–	–	–	

Enzymes for which no activity was measurable ($k_{cat} < 0.001\ s^{-1}$) are denoted with dashes.

Table 4.3: Kinetic parameters for 2-deoxyR5P as an alternate aldose substrate to A5P for wild-type *Nme*KDO8PS and *Afe*KDO8PS and mutants *Nme*N59A and *Afe*N57A.

KDO8PS	2-deoxyR5P			Ratio of specificity constants
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\,mM^{-1}$)	A5P/2-deoxyR5P
*Nme*WT	0.230 ± 0.020	0.13 ± 0.01	0.57 ± 0.09	1200
*Nme*N59A	0.432 ± 0.037	0.0070 ± 0.0002	0.016 ± 0.002	18
*Afe*WT	5.0 ± 0.8	0.66 ± 0.06	0.13 ± 0.03	1700
*Afe*N57A	1.5 ± 0.1	0.202 ± 0.006	0.13 ± 0.01	28

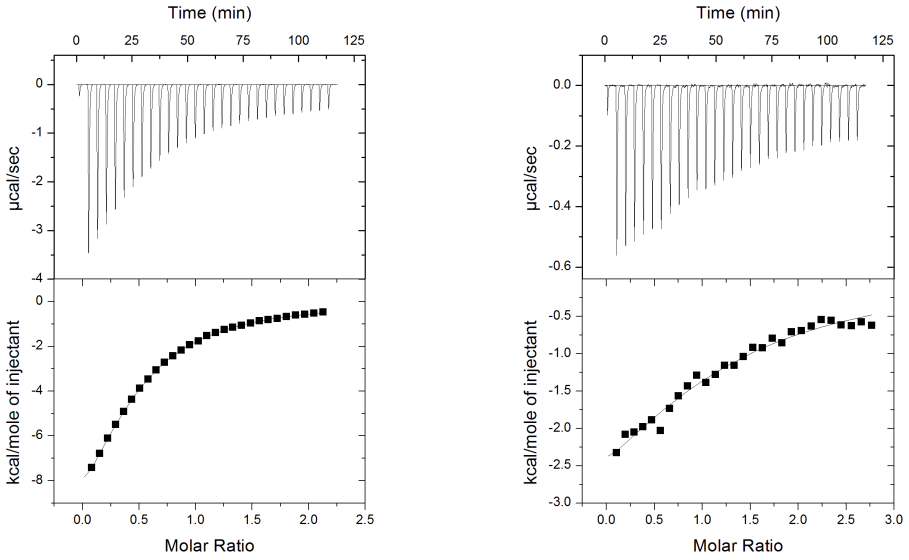
Indeed, for *Afe*KDO8PS, 2dR5P is as good a substrate for the N57A mutant as it is for the wild-type enzyme. This observation is completely consistent with an important role of this Asn in the KANRS motif for ensuring selection by these enzymes of the substrate with the correct stereochemistry (Figure 4.2).

4.4 PEP binding is altered by mutation of the conserved Lys of the KANRS motif

The Lys of the KANRS motif lies close to the PEP-carboxylate binding site and may play a role in the binding of PEP; in *Aae*KDO8PS, this residue is observed to hydrogen bond to PEP, A5P, or both.[24] To assess the effect of the mutation of this residue, PEP binding for both *Nme*K57A and *Afe*K55A was studied by ITC. Weak binding of PEP to *Nme*K57A was observed, with a K_{d} value of 56 ± 14 µM (Figure 4.4b). Binding of a similar weak nature was observed for PEP with *Afe*K55A [K_{d} was 57.5 ± 0.6 µM (Figure 4.4a)]. These values are 10 times greater than those calculated when the same experiment was performed with the wild-type protein (see Chapter 3, K_{d}^{Nme} was 4.8 ± 0.2 µM; K_{d}^{Afe} was 5.0 ± 0.1 µM), consistent with significant disruption to PEP binding upon mutation of this absolutely conserved Lys residue of the KANRS motif.

4.5 Crystallography

The crystal structure of wild-type *Nme*KDO8PS has previously been determined,[39] and similarly, the structures of *Nme*K57A, *Nme*N59A (crystallised and solved by Dr Fiona Cochrane and Professor Geoff Jameson, Massey University), *Nme*KARS, and *Nme*KPRS feature one tetramer in each asymmetric unit. As expected, the construction of each subunit is unchanged in a $(\beta/\alpha)_8$ TIM-barrel topology. As in the wild-type structure, loops β7α7 and β8α8 are



(a) **(b)**

Figure 4.4: Interaction of PEP with (a) *Afe*K55A and (b) *Nme*K57A, quantified by ITC.

disordered, showing non-continuous electron density. Loop β2α2, on which the KANRS motif and these mutations are located, is well-defined in chains B and C and is moderately to poorly defined, in parts, in chains A and D. The *Nme*KARS mutant, in particular, shows non-continuous electron density in these loops. Notwithstanding the lower resolution of the *Nme*KPRS data set, electron-density maps were remarkably clean with the main chain and proline of the mutated region in the *Nme*KPRS mutant clearly defined and with the arginine side chain of the KPRS motif well-defined in three of the four chains.

For both *Nme*K57A and *Nme*N59A, the mutations are clearly visible. The only apparent change in the crystal structures is the respective truncation of the amino acid side chain. Few other changes are observed in the structures (Figure 4.5), especially for the *Nme*N59A mutant, which superimposes very closely on the wild-type structure with a RMSD of 0.161 Å for superposition of Cα atoms of the tetrameric assembly. For the *Nme*K57A mutant, the

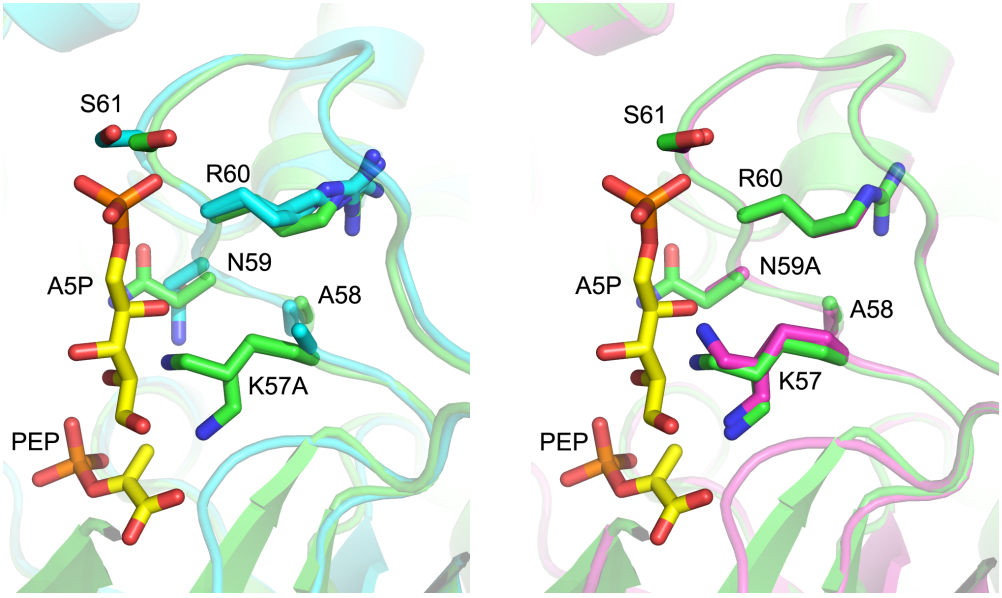
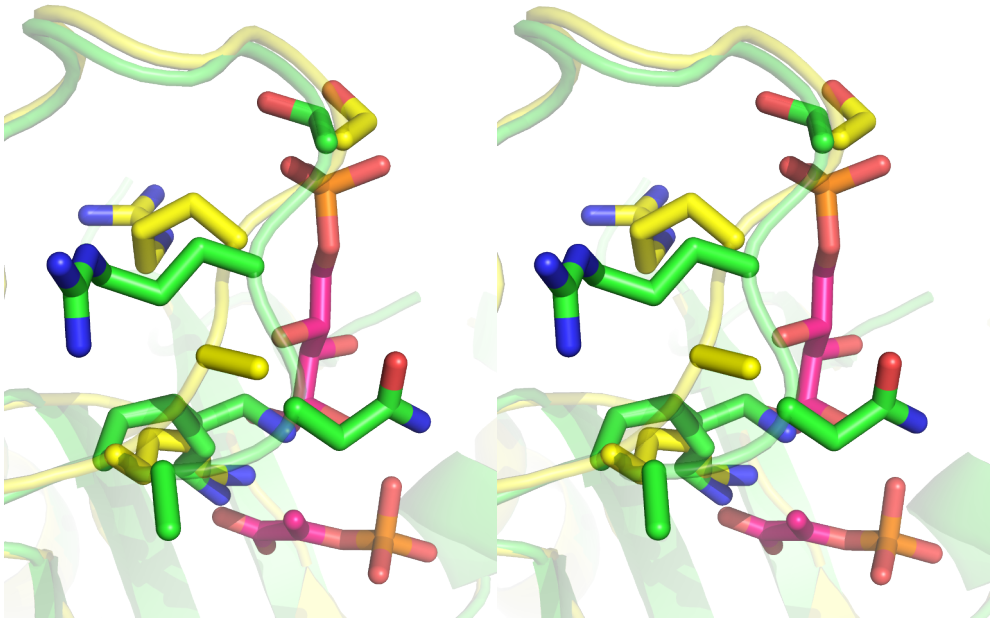


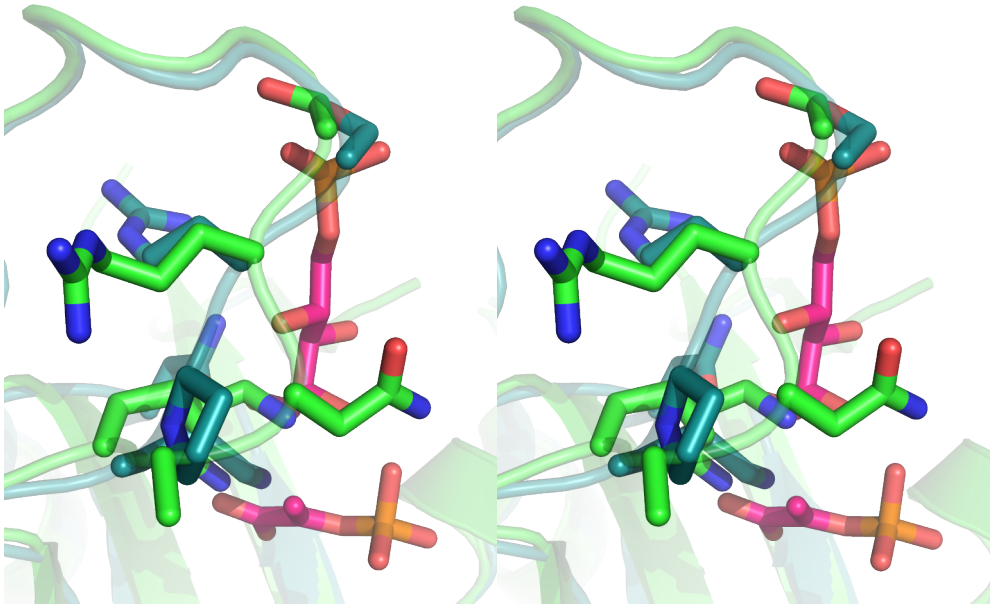
Figure 4.5: Superposition of the active sites of wild-type *Nme*KDO8PS (green), *Nme*K57A (cyan), and *Nme*N59A (magenta). (a) Wild type and *Nme*K57A. (b) Wild type and *Nme*N59A. Shown is chain C of each crystal structure. PEP and A5P are shown with carbon atoms coloured yellow and from a superposition of the structure of *Aae*KDO8PS (PDB code 2NX3). Lys57 is modelled in two alternative conformations in both wild-type *Nme*KDO8PS and *Nme*N59A. Arg60 in *Nme*K57A is also modelled in two alternative conformations.

corresponding RMSD is substantially greater (0.388 Å); the loss of the active-site lysine side chain leads to a slight repositioning of the β2α2 loop bearing the K57A mutation in chains A, B, and D, as well as in chain D repositioning of the β1α1 loop bearing the active-site Asn23 (to which Lys57 hydrogen bonds in the wild-type, *Nme*N59A, *Nme*KARS, and *Nme*KPRS structures).

In the structure of *Nme*KARS (RMSD for Cα atoms of 0.314 Å for superposition onto the wild-type structure), the shortening of the motif causes the backbone trace of the protein to take a shortcut (Figure 4.6a). In chain A, the main chain shift is from Arg59 to Ile62 (from Arg60 to Ile63 in the wild type, respectively); the Cα of Arg59 is positioned approximately where that of Asn59 of the wild-type enzyme resided, and the side chain of Arg59 extends in the same direction as (and beyond that of) Asn59 to hydrogen



(a)



(b)

Figure 4.6: (a) Stereoview showing a superposition of chain C of *Nme*KARS (yellow) and wild-type *Nme*KDO8PS (green). (b) Stereoview showing a superposition of chain C of *Nme*KPRS (teal) and wild-type *Nme*KDO8PS (green). To indicate the substrate binding sites, PEP and A5P (carbon atoms coloured magenta) are shown from a superposition of the structure of *Aae*KDO8PS (PDB code 2NX3). The residues belonging to the native or mutated motif are shown as sticks.

bond with Gln117 of the adjacent subunit C. Ser60 and Ser61 occupy the positions of Arg60 and Ser61 in the wild-type structure, respectively. Register with the corresponding chain of the wild-type structure is not re-established until Ile62 (Ile63 in the wild type). The remainder of the β2α2 loop is not observed in chain A (residues 63–68). In chains B and C, the main chain is shifted between Lys57 and Arg59 (exclusive), meaning that Ser60 and Ser61 are in positions similar to those of the wild-type enzyme (Ser61 and Ser62, respectively). The Cα atom of Arg59 in subunits B and C is positioned approximately in place of that of Arg60 of the wild-type enzyme. However, unlike wild-type *Nme*KDO8PS where Arg60 projects into the interface with subunits D and A, forming an intersubunit salt bridge with Asp120 on loop β4α4, in *Nme*KARS the side chain is oriented like that seen in the subunit of *Aae*KDO8PS that binds A5P (PDB code 1FWW). Chain D in this region is completely disordered, representing in extreme form the conformational plasticity of the β2α2 loop.

In the structure of *Nme*KPRS (Figure 4.6b), the backbone trace and the side-chain conformations are similarly oriented and are very similar to those of *Nme*KARS, with the obvious exception of Pro58. In contrast to *Nme*KARS, the main chains of subunits A and D are defined for the β2α2 loop, except that the side chain for Arg59 of chain A is not defined in electron-density maps and Arg59 in chain D follows a trajectory similar to that in chain A of *Nme*KARS. The RMSD of the Cα atoms for superposition onto the wild-type structure is 0.294 Å.

Table 4.4: Crystal parameters, data collection, and refinement statistics for *Nme*K57A and *Nme*N59A.

	*Nme*K57A	*Nme*N59A
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	82.43, 86.18, 163.31	81.64, 85.37, 162.62
Resolution range (Å)	34.59–1.95 (2.02–1.95)	39.93–1.75 (1.81–1.75)
No. of measurements	431450	424611
No. of unique reflections	83787	106837
Redundancy	5.15 (5.16)	3.97 (2.98)
Completeness (%)	98.0 (100)	92.8 (65.2)
$I/\sigma(I)$	4.2 (1.6)	8.8 (1.9)
R_{merge}	0.079 (0.394)	0.036 (0.355)
Wilson B value (Å^2)	33.8	34.1
	Refinement	
Resolution (Å)	33.84–1.95 (2.00–1.95)	40.00–1.75 (1.80–1.75)
R_{cryst}	0.2048	0.1800
R_{free}	0.2521	0.2152
Amino acids (chain length of 280 residues)	250 + 250 + 251 + 249 residues; 7827 atom sites	251 + 252 + 252 + 254 residues; 7828 atom sites
No. of water molecules	737	698
No. of others	3 (1×0.5) Cl^-, 1 Na^+, 1 glycerol	8 Cl^-, 1 Na^+, 0.5 PO_4^{2-}, 2 (1×0.5) glycerol
Mean B (Å^2)		
Protein	36.44	31.53
Water	43.43	39.85
Other	25.44	38.75
RMSD from target values		
Bond lengths (Å)	0.010	0.010
Bond angles (°)	1.090	1.201
Dihedral angles (°)	6.623	5.402
Ramachandran		
Most favoured (%)	92.4	93.5
Allowed (%)	6.6	5.5
Generously allowed (%)	0.6	0.7
Disallowed (%)	0.5	0.3
PDB code	3QPY	3QPZ

Table 4.5: Crystal parameters, data collection, and refinement statistics for *Nme*KARS and *Nme*KPRS.

	*Nme*KARS	*Nme*KPRS
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	82.03, 85.81, 163.14	81.82, 85.15, 163.78
Resolution range (Å)	36.84–1.90 (1.97–1.90)	55.47–2.70 (2.85–2.70)
No. of measurements	412547	150576
No. of unique reflections	91249	32102
Redundancy	4.52 (4.36)	4.7 (4.80)
Completeness (%)	99.4 (97.3)	99.7 (99.5)
$I/\sigma(I)$	5.1 (2.1)	5.9 (1.5)
R_{merge}	0.070 (0.357)	0.119 (0.503)
Wilson B value (Å^2)	26.5	45.8
	Refinement	
Resolution (Å)	36.66–1.90 (1.95–1.90)	47.89–2.70 (2.77–2.70)
R_{cryst}	0.1887	0.2112
R_{free}	0.2285	0.2749
Amino acids (chain length of 280 residues)	243 + 247 + 249 + 240 residues; 7532 atom sites	249 + 251 + 249 + 251 residues; 7711 atom sites
No. of water molecules	717	139
No. of others	4 Cl^-, 2 Na^+, 2 glycerol	none
Mean B (Å^2)		
Protein	28.19	34.35
Water	36.60	32.77
Other	30.75	–
RMSD from target values		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	0.932	1.115
Dihedral angles (°)	6.437	5.082
Ramachandran		
Most favoured (%)	92.6	93.2
Allowed (%)	6.2	5.7
Generously allowed (%)	0.6	0.6
Disallowed (%)	0.6	0.5
PDB code	3QQ0	3QQ1

4.6 Discussion

Mutants of the conserved active-site KANRS motif in *Nme*KDO8PS and *Afe*KDO8PS were successfully created, expressed, and purified. The purified proteins exhibited stability characteristics and overall structure similar to those of the wild-type proteins. The mutations did, however, affect the enzymatic functioning. The notable feature of these analyses is that any mutation of this motif significantly prohibited both metal-dependent and metal-independent enzymes from playing their normal catalytic role.

4.6.1 Lys is essential for catalysis

Substitution of the KANRS Lys with Ala was completely deleterious for enzyme activity, and the role of this residue has never previously been directly probed in KDO8PS, despite being absolutely conserved and present in a key active-site location. The most recent structural analysis of both metal-dependent and metal-independent enzymes suggests that the amino functionality of this Lys is in the proximity of the A5P aldehydic functionality. In this position, the aldehyde group is held in place by this Lys residue, and by coordination to the metal ion for metal-dependent KDO8PSs, or by the Cys-substituted Asn for metal-independent KDO8PSs.[29] This Lys clearly plays a role in the correct binding of A5P. The structure of the *Nme*KDO8PS K57A mutant suggests that the loss of the lysine side chain results in no significant structural change. Computational studies have suggested that the KDO8PS reaction proceeds by the olefinic atom C3 of PEP attacking the aldehydic carbonyl.[19,29] Protonation of the aldehydic carbonyl is required, and it appears likely that this Lys performs this critical catalytic role (Figure 4.7), either directly or via the nearby Asp residue (Asp233 in *Aae*KDO8PS) that has also been shown to be essential.[39] Following attack by PEP on C1 of A5P, water is known to add to C2 of PEP, generating a linear tetrahedral diphosphate intermediate.[22] Computational studies have favoured this water to approach from the *si* face of PEP giving an overall syn addition of A5P

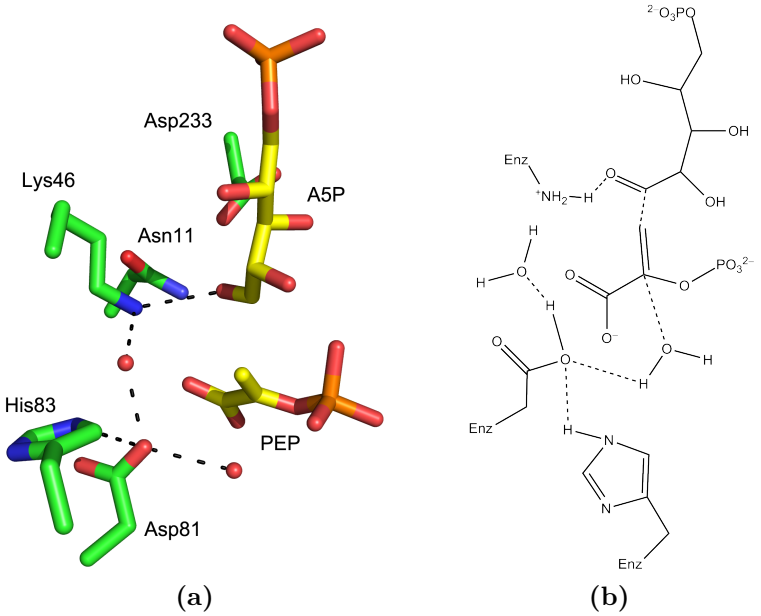


Figure 4.7: The proton relay chain that may link deprotonation of the nucleophilic water molecule on the *re* face of PEP with protonation of the carbonyl oxygen of A5P. (a) Crystal structure model of *Aae*KDO8PS (PDB code 2NX3) showing the interactions from the *re*-face water molecule to the A5P carbonyl. The carbon atoms of PEP and A5P are coloured yellow. (b) Possible interactions involved in catalysis.

and water across the double bond of PEP.[19,29] However, it is apparent that the essential Lys of the KANRS motif links (via an absolutely conserved Asp, Asp92 in *Nme*KDO8PS and Asp81 in *Aae*KDO8PS) to a water molecule on the *re* face of PEP. This water molecule is observed and crystallographically conserved in most KDO8PS structures and is located ≈2.8 Å from the olefinic moiety of PEP. This arrangement, and the importance of this Lys revealed by these studies, suggest that this *re*-side water perhaps warrants further consideration, and that this residue may play a dual role in A5P binding and protonation. Attack by water in this arrangement would be in line with the proposed mechanisms for DAH7PS.[74,76]

4.6.2 Asn has a role in substrate selection

In contrast to mutation of the Lys, mutation of the Asn residue of the KANRS motif to Ala did produce enzymes with some limited catalytic function. As the structure of the N59A mutant of *Nme*KDO8PS revealed no changes to the active site, this alteration in catalytic function is likely to be due to the loss of the amide side-chain functionality. These enzymes were substantially compromised in their ability to bind A5P, but intriguingly were able to utilise 2dR5P, which is a poor alternative substrate for the wild-type enzymes, relatively well. For *Afe*KDO8PS, the N57A mutant was shown also to tolerate 2dR5P in a manner similar to that of the wild-type enzyme.

Structures of substrate-bound *Aae*KDO8PS show the Asn residue of the KANRS motif hydrogen-bonded to the C2 hydroxyl group of A5P (Figure 4.2a). The C2 hydroxyl group also forms a hydrogen bond with the phosphate group of PEP in the *Aae*KDO8PS structure. The importance of these interactions for the reaction catalysed by wild-type KDO8PSs is revealed clearly by the observations that R5P with an altered C2 configuration is unable to act as an alternative substrate for the enzymes, and that 2dR5P (equivalent to 2-deoxyA5P) is an extremely poor substrate. The importance of the interaction between the C2 hydroxyl group of A5P and the wild-type enzyme is entirely consistent with the low activity of the Asn-to-Ala mutants (N59A and N57A), and the relative ambivalence of these mutants to the presence of the C2 hydroxyl group of A5P. Indeed, the low activity of both of these mutants mirrors, in part, the effects observed for the wild-type enzymes when the C2 hydroxyl group of A5P is removed.

Interestingly, the metal-dependent enzymes were found to be far more tolerant of changes to the aldehydic substrate. 2dR5P is a slightly better substrate for wild-type *Afe*KDO8PS than for the metal-independent *Nme*-KDO8PS, and the N57A *Afe*KDO8PS mutant is slightly less compromised than the corresponding *Nme*KDO8PS mutant and notably demonstrates the same ability to utilise 2dR5P as wild-type *Afe*KDO8PS. This suggests that metal coordination plays a more prominent role in competent substrate

binding than the equivalent amide functionality of the Asn residue that substitutes for the metal-binding Cys ligand to give the metal-independent enzymes. Mutation of Asn of the KANRS motif to Asp was lethal to activity. This observation is consistent with the inability of Asp, through the absence of a hydrogen-bond donor, to perform a function similar to that of the Asn in this position.

Intriguingly, DAH7PS enzymes are far more tolerant of changes at C2 of the natural substrate E4P, of which the C2 configuration is the opposite to that of A5P.[59] E4P is shorter with fewer degrees of freedom and therefore may not require the same anchoring that the A5P substrate requires to achieve reaction. DAH7PSs are also all metal dependent, and metal coordination may provide the key interaction that allows competent placement of E4P for these enzymes.

4.6.3 KARS and KPRS mutants lose activity

Given the huge reduction in catalytic power upon the loss of the Asn of the KANRS motif, it is not surprising that the KARS and KPRS mutations exhibited no KDO8PS activity. As well as missing the key substrate-binding contact with the Asn of the KANRS motif, as confirmed by the *Nme*KDO8PS structures of these mutants, the substrate-binding site has been shortened with the key phosphate-binding Arg of the KANRS motif occupying a position similar to that of Asn59 in wild-type *Nme*KDO8PS. However, the change from KANRS to KPRS was not sufficient to re-create DAH7PS activity, as these enzymes also did not utilise E4P. These results indicate very clearly that other features of the protein are important for supporting substrate selection and determining productive aldose phosphate binding.

While the KPRS motif is thought to be important for determining E4P selection in DAH7PSs, analysis of both the sequences and structures of these enzymes indicates some of the features of KDO8PS that may preclude these *Nme*KDO8PS and *Afe*KDO8PS mutants from utilising E4P. There are other notable features that also may have a bearing on substrate selection

in conjunction with the KPRS motif. The E4P/A5P binding site is situated close to the interface region between adjacent monomer units in the tetrameric enzyme. There are no interdigitating or shared residues between adjacent monomers in DAH7PS that might contribute directly to the active site. However, for KDO8PS, an Arg (R117 in *Nme*KDO8PS, R115 in *Afe*KDO8PS, or R106 in *Aae*KDO8PS, which is part of a conserved PAFLxR motif in KDO-8PS sequences) also makes contacts with the A5P phosphate. This Arg has been investigated previously in *Aae*KDO8PS and shown to be important for KDO8PS catalysis.[79] This residue provides a salt bridge to the A5P phosphate and facilitates closure of the β7α7 loop (through precisely positioning the KANRS-containing β2α2 loop). This Arg is still present in both the KPRS mutants, and its presence may discourage correct E4P binding. In contrast, DAH7PSs may provide Arg functionality to the aldose phosphate binding site from within the monomer unit, from the equivalent GARNxQ sequence motif. The β7α7 loop that is closed on A5P binding is another obvious sequence and structural difference between DAH7PS and KDO8PS. Assisted by the intersubunit Arg, this loop closes upon A5P binding.[79] Sequence alignments reveal that this β7α7 loop is absent in DAH7PS. These differences between KDO8PS and DAH7PS are investigated in the next chapters.

4.7 Conclusion

In summary, the studies in this chapter have shown that the KANRS motif is of prime importance to the function of KDO8PS. The Lys is critical, and this residue may act as a general acid in the reaction mechanism, thereby activating A5P directly to attack by cosubstrate PEP. The Asn of this motif appears to play an important role in the selection of the correct substrate by the enzyme. Simple conversion of this motif into the KPRS motif found absolutely conserved in DAH7PSs is not sufficient to switch substrate specificity from A5P to E4P, but this conversion is catastrophic for KDO8PS activity.

Chapter 5

The role of the KDO8PS-unique β7α7 loop

5.1 Introduction

In earlier chapters, the roles of the residues on the β2α2 and β8α8 loops in KDO8PS have been examined and discussed. The β7α7 loop is the third loop that extends from the C-terminal end of the barrel in KDO8PS and completes the A5P binding site of the enzyme active site (Figure 5.1a). A notable difference between KDO8PSs and DAH7PSs is highlighted by sequence and structural alignments that indicate this long or extended β7α7 loop is always present in KDO8PSs, yet is not found in DAH7PSs (Figure 5.1 and Figure 5.2a).[79] This substantial sequence difference between the two enzymes represents a clear conserved difference in the construction of the binding sites for A5P and E4P. In KDO8PSs, this β7α7 loop is defined in crystal structures only when A5P and PEP are bound. Additionally, under these conditions, only two of the four active sites in the tetrameric assembly are observed to bind A5P simultaneously, suggesting that in KDO8PS catalysis may take place at alternating active sites.[72]

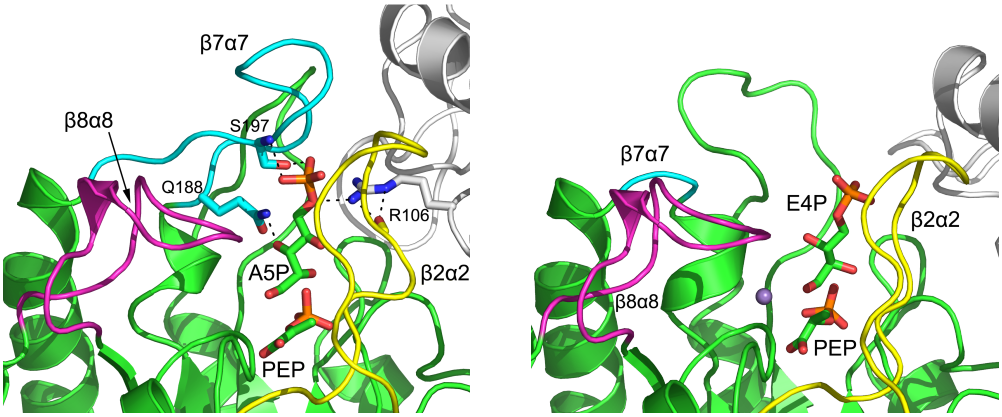


Figure 5.1: Structures of KDO8PS and DAH7PS showing the loops that extend from the core barrel and contribute to the active-site architecture. (a) *Aae*KDO8PS (PDB code 2NX3). The conserved residues on the β7α7 loop are shown in stick form, as is the interdigitating Arg106 from the adjacent subunit. (b) *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]). In both structures, the β7α7 loop is coloured cyan, the β8α8 loop is coloured magenta, and the β2α2 loop is coloured yellow.

Initially, it was proposed that the β7α7 loop played a role in shielding the active site from bulk solvent, to prevent the undesirable reaction of PEP with water.[80] Recently, however, it was found that mutation of a conserved Arg on loop β4α4, which interacts with the β2α2 loop across a subunit interface, and contributes to the binding of the A5P phosphate moiety (Figure 5.1a), impaired closure of the β7α7 loop, without leading to the unproductive reaction of PEP with water.[79] The role of this conserved difference in the β7α7 loop between KDO8PS and DAH7PS remains unclear.

In this chapter, the role of the β7α7 loop in KDO8PS activity is directly examined, via its removal to match the active-site architecture of DAH7PS and via the mutation of the two absolutely conserved residues of this loop, a Gln and a Ser (Figure 5.1). The results indicate that the extent of catalysis by both metal-dependent and metal-independent enzymes is significantly diminished by the single mutations and combination of the two single mutations, and that catalysis is almost completely abrogated in the absence of this loop. The multiple interactions provided by this extended loop are required to support

efficient catalysis by KDO8PS.

5.2 Preparation of β7α7 loop mutants

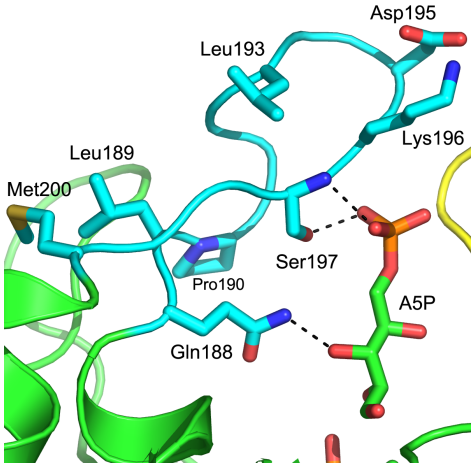
Relative to the short loop linking β7 and α7 in DAH7PS, the β7α7 loop of both metal-dependent and metal-independent KDO8PSs is extended by 11 residues (Figure 5.2). Two of these residues, Gln and Ser, at the beginning and near the end of the loop, respectively, are absolutely conserved in KDO8PSs. The structure of the *Aae*KDO8PS enzyme in complex with PEP and A5P indicates contacts are made between the Gln and C3-OH group of A5P and between the Ser and the A5P phosphate functionality (Figure 5.2b).

To probe the role of the conserved Gln and Ser residues in the extended β7α7 loop of KDO8PS, these residues were mutated to Ala in both the metal-dependent *Afe*KDO8PS and metal-independent *Nme*KDO8PS (creating *Afe*Q198A, *Afe*S207A, *Nme*Q202A and *Nme*S211A). The mutagenesis to create these single mutant constructs (except *Afe*Q198A) was performed by Parker research group members Ben Gloyne and Dr Richard Hutton.[95] Double-mutant proteins containing both the Gln-to-Ala and Ser-to-Ala substitutions (*Nme*Q202A/S211A and *Afe*Q198A/S207A) were also created. Additionally, the 11 extra residues of this extended loop were excised from both proteins to make them resemble this loop's length in DAH7PS, creating *Nme*L7trun and *Afe*L7trun. The constructs containing only the β7α7 loop excisions were created by Parker research group members Evan Nimmo and Dr Richard Hutton.[96] These truncated enzymes were further modified in the β2α2 loop region, via alteration of the absolutely conserved LysAlaAsnArgSer (KANRS) motif of KDO8PS to the equivalent LysProArgSer (KPRS) motif found for DAH7PSs, creating *Afe*KPRS/L7trun and *Nme*KPRS/L7trun.

The mutants were expressed and purified using the same methods that were used for the respective wild-type proteins.[39,84] Measurement of the CD spectra of the mutated *Afe*KDO8PS proteins indicated that the introduced mutations had caused no gross structural perturbations (Figure 5.3). Denat-

```
AaeKDO8PS    AKVIYDATHSVQLPGGLGDKSGGMREFIFPL
HpyKDO8PS    APVIFDATHSVQMPGGANGKSSGDSSFAPIL
AfeKDO8PS    CPVVFDATHSVQLPGGQGDRSGGQREFIPVL
NmeKDO8PS    LPVIFDVTHSLQTRDAGSAASGGRRAQALDL
EcoKDO8PS    SPVIFDVTHALQCRDPFGAASGGRRAQVAEL
TmaDAH7PS    LPILVDPSHSG----------GRRDLVIPL
PfuDAH7PS    LPIIVDPSHPA----------GRRSLVIPL
ApeDAH7PS    LPVIVDPSHPA----------GRRSLVPAL
SpyDAH7PS    LPIIVDVSHST----------GRRDLLLPA
               :: * :*           . *
```

(a)



(b)

Figure 5.2: (a) Partial sequence alignment of KDO8PS and type Iβ DAH7PS sequences showing the additional amino acids that extend the β7α7 loop in KDO-8PSs. The regions corresponding to those excised in *Nme*KDO8PS and *Afe*KDO8PS are highlighted in yellow, inclusive of the two conserved residues of the β7α7 loop, which are highlighted in cyan. *Hpy*: *H. pylori*, *Tma*: *T. maritima*, *Ape*: *A. pernix*, *Spy*: *S. pyogenes*. (b) Expansion of the β7α7 loop of *Aae*KDO8PS in Figure 5.1a. In this chapter, the absolutely conserved Gln and Ser of the β7α7 loop (equivalent to Gln188 and Ser197 of *Aae*KDO8PS, respectively) were mutated to Ala in *Nme*-KDO8PS and *Afe*KDO8PS, creating *Nme*Q202A, *Nme*S211A, *Nme*Q202A/S211A, *Afe*Q198A, *Afe*S207A, and *Afe*Q198A/S211A.

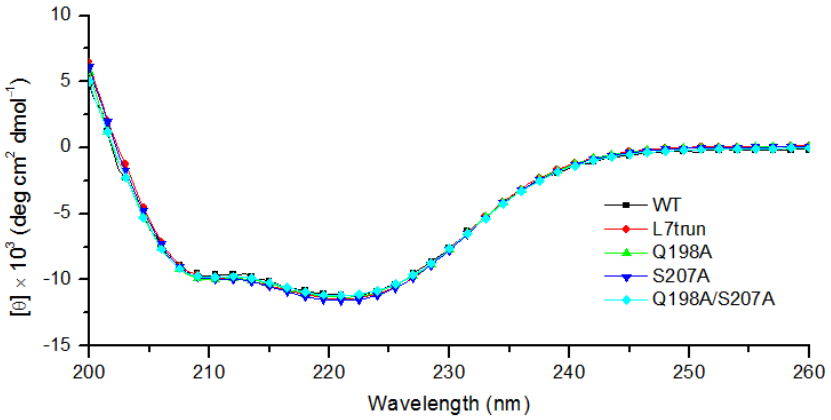


Figure 5.3: CD spectrophotometry of *Afe*KDO8PS wild-type and β7α7 loop mutants. Wild-type: black squares; *Afe*L7trun: red circles; *Afe*Q198A: green triangles; *Afe*S207A: blue triangles; *Afe*Q198A/S207A: cyan diamonds. Every third datum point is plotted.

uration temperatures measured by DSF for the β7α7 loop-truncated proteins of both *Nme*KDO8PS and *Afe*KDO8PS (Table 5.1) were also measured. The denaturation temperatures for the *Nme*KDO8PS β7α7 loop truncated protein were similar to those of the wild type, and the presence of Cd^{2+}, which results in a substantial reduction in the denaturation temperature of the wild-type enzyme, remained strongly destabilising. The *Afe*KDO8PS β7α7 loop-truncated protein was less stable under all conditions, showing a denaturation temperature lower than that of wild-type *Afe*KDO8PS.

5.3 Kinetic characterisation

The single- and double-mutant proteins and full loop truncation proteins (*Nme*L7trun and *Afe*L7trun) all retained some ability to catalyse the KDO8PS reaction (Table 5.2). However, the proteins in which the β7α7 loop had been removed catalysed the KDO8PS reaction very poorly. For all active mutant enzymes, $K_{\mathrm{m}}^{\mathrm{PEP}}$ was similar to that measured for the respective wild-type

Table 5.1: The effect of additives on T_m for β7α7 loop truncated mutants of *Nme*KDO8PS and *Afe*KDO8PS.

KDO8PS	Additives	T_m (°C)	Diff.[a]	Diff. to WT
*Nme*L7trun	No additive	57.0 ± 0.3		−1
	Mn^{2+}	57.0 ± 0.3	0	−2
	Cd^{2+}	43 ± 1	−14	1
	PEP	59 ± 1	2	−1
*Afe*L7trun	No additive	61 ± 1		−7
	Mn^{2+}	61 ± 1	0	−7
	Cd^{2+}	70 ± 1	9	−12
	PEP	61 ± 1	0	−8

[a] Differences are with respect to no additive conditions.

Table 5.2: Kinetic parameters for wild-type and β7α7 loop mutants of *Nme*-KDO8PS and *Afe*KDO8PS.

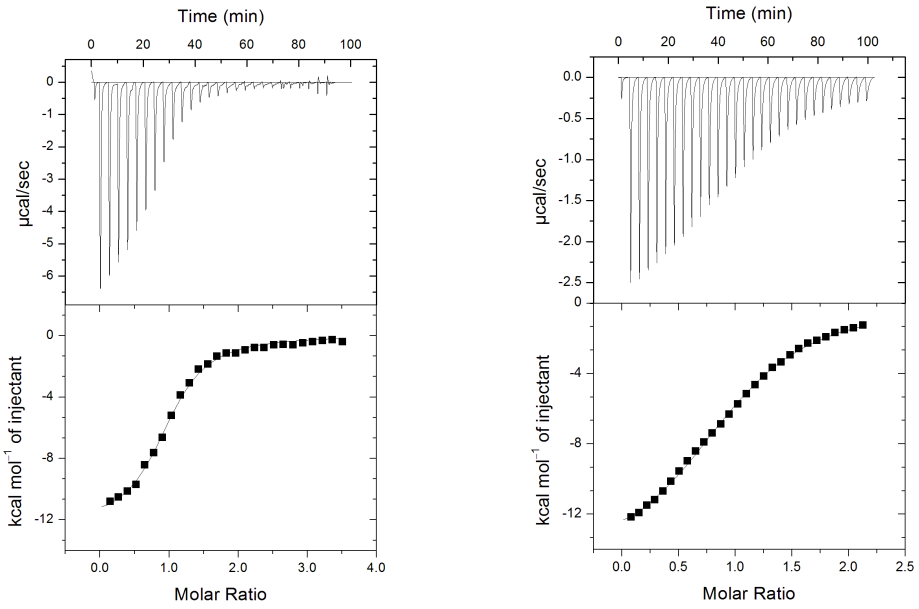
KDO8PS	K_m^{PEP}	K_m^{A5P}	k_{cat}	k_{cat}/K_m^{A5P}
*Nme*WT	2.5 ± 0.2	12.0 ± 0.5	8.0 ± 0.1	670 ± 40
*Nme*Q202A	7 ± 1	160 ± 20	3.6 ± 0.2	23 ± 4
*Nme*S211A	3.6 ± 0.9	74 ± 10	6.2 ± 0.3	80 ± 20
*Nme*Q202A/S211A	9 ± 1	300 ± 18	3.8 ± 0.2	13 ± 1
*Nme*L7trun	[a]	3900 ± 100	0.46 ± 0.01	0.118 ± 0.007
*Nme*KPRS/L7trun	–	–	–	–
*Afe*WT	12 ± 1	22 ± 2	4.8 ± 0.1	220 ± 20
*Afe*Q198A	11 ± 1	151 ± 13	3.1 ± 0.2	21 ± 3
*Afe*S207A	6.8 ± 0.3	153 ± 17	2.53 ± 0.03	17 ± 2
*Afe*Q198A/S207A	59 ± 3	173 ± 23	1.38 ± 0.03	8 ± 1
*Afe*L7trun	[a]	480 ± 20	0.0728 ± 0.0008	0.151 ± 0.007
*Afe*KPRS/L7trun	–	–	–	–

K_m values are in units of µM, k_{cat} in s^{-1}, and k_{cat}/K_m in $s^{-1}\,mM^{-1}$. Enzymes with no measurable activity ($k_{cat} < 0.001\ s^{-1}$) are denoted with dashes.
[a] Not measured

enzyme. In contrast, the $K_{\mathrm{m}}^{\mathrm{A5P}}$ values were higher than those recorded for the wild type, consistent with the mutated regions of the proteins making weaker interactions with this substrate. For *Afe*KDO8PS, both single and double substitutions with Ala of the conserved Gln and Ser of the β7α7 loop resulted in similar ≈7- and 8-fold increases in $K_{\mathrm{m}}^{\mathrm{A5P}}$, respectively, relative to that of the wild-type enzyme. On the other hand, for *Nme*KDO8PS, there are 6-, 12-, and 24-fold increases in $K_{\mathrm{m}}^{\mathrm{A5P}}$ for mutants S211A, Q202A, and Q202A/S211A, respectively, relative to that of the wild-type enzyme, suggesting that in this enzyme the binding interactions provided by the Ser to A5P are less important that those provided by the Gln.

In marked contrast to the modest loss of catalytic activity of the single and double mutants, the truncated proteins (*Nme*L7trun and *Afe*L7trun), without the extended β7α7 loop, exhibited very low activity. $K_{\mathrm{m}}^{\mathrm{A5P}}$ values were much larger, and the k_{cat} values for these enzymes were significantly lower. For the metal-independent *Nme*KDO8PS enzyme, the change in specific activity was more substantial: the $k_{\mathrm{cat}}/K_{\mathrm{m}}^{\mathrm{A5P}}$ was found to be more than 5700 times lower for the truncated protein than for the wild-type enzyme (compared to 1400 times lower for the metal-dependent *Afe*KDO8PS). There was no measurable activity for the proteins that combined the β7α7 loop truncation with the mutation of the conserved KANRS motif, to the equivalent KPRS motif conserved in DAH7PS (KPRS/L7trun). Nor was there activity for these KPRS/L7trun mutated enzymes with R5P, 2dR5P, or E4P used in place of A5P.

The binding interactions of PEP with the β7α7 loop-truncated proteins were measured by ITC. Compared to measurements on the wild-type protein (for *Afe*KDO8PS, K_{D} is 5.0 ± 0.1 µM; for *Nme*KDO8PS, K_{D} is 4.8 ± 0.2 µM), the PEP binding to the truncated proteins is significantly but not greatly impaired (for *Nme*L7trun, K_{D} is 11.8 ± 0.8 µM; for *Afe*L7trun, as measured by Parker research group members Evan Nimmo and Dr Richard Hutton,[96] K_{D} is 13.6 ± 0.4 µM) (Figure 5.4), indicating that the primary effect of the deletion of the β7α7 loop was to impair the interaction with A5P.

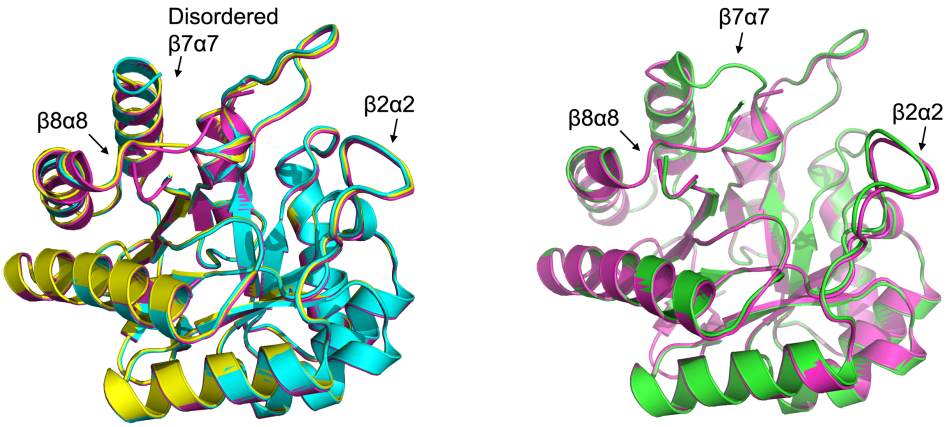


(a) **(b)**

Figure 5.4: The interaction between the truncated β7α7-loop mutants and PEP, quantified by ITC. (a) *Afe*L7trun (measured by Evan Nimmo and Dr Richard Hutton[96]). (b) *Nme*L7trun.

5.4 Crystallography

The *Nme*KDO8PS mutant proteins crystallised in the same way as the wild-type enzyme with one complete tetramer per asymmetric unit. In the structure of wild-type *Nme*KDO8PS (and those of its mutants), the residues of the β7α7 loop are disordered from Gln202 (or Thr203) to at least Gly213 in all four subunits (Figure 5.5a). In the structure of *Nme*Q202A (RMSD of Cα atoms on the superimposed wild-type structure of 0.161 Å), there are no observed differences compared to the wild-type structure. The only exception is for the Q202A mutation, which, as in the wild-type structure, is the last ordered residue of the β7α7 loop. Likewise, the structure of *Nme*S211A is very similar to the wild-type structure (RMSD of 0.233 Å). However, for this mutant, loop β2α2 of chain D has shifted away from the active site between residues Asn59



(a) **(b)**

Figure 5.5: (a) Superposition of the structures of *Nme*Q202A (yellow), *Nme*S211A (cyan), and wild-type *Nme*KDO8PS (magenta). (b) Superposition of the structures of *Nme*L7trun (green) and wild-type *Nme*KDO8PS (magenta). The β8α8 loop is partially ordered in all structures, and the β7α7 loop is ordered only in *Nme*L7trun. Subunit C has been used in each superposition.

and Val69, and in chain B, the side chain of Glu236 now reaches into the space usually occupied by His199 and that of His199 has moved to a space sometimes occupied by alternative conformations of Arg165. The mutation of Ser211 to Ala is in the region of loop β7α7 that is disordered in the wild-type and all non-truncated mutant structures.

The structure of *Nme*L7trun is also very similar to the wild-type structure, but the RMSD of 0.401 Å, while modest, is significantly greater than that for superposition of single and double mutants on the wild-type structure, indicating that there are small changes induced by the truncation. The most notable of these changes is a small twist at the end of α-helix 7, which is likely constrained relative to the wild-type protein because of the shortened β7α7 loop (Figure 5.5b). However, for the first time, except for KDO8PSs from hyperthermophilic organisms, the truncated β7α7 loop is ordered in all four subunits. In chain B, as was observed for *Nme*S211A, the side chains of His199, Glu228, and Arg165 have likewise rearranged, suggesting that in the absence of substrates there is some flexibility in the positioning of these side

chains.

The *Nme*KPRS/L7trun mutant protein crystallised in the same space group as all previous *Nme*KDO8PS structures but with different unit-cell dimensions. The crystal structure is very similar to that of *Nme*L7trun (RMSD of 0.401 Å). In all *Nme*KDO8PS structures, both the (non-truncated) β7α7 and β8α8 loops are normally disordered. For the first time in a structure of *Nme*KDO8PS, the main chain of the β8α8 loop has become ordered (in chain B) (Figure 5.6). This is likely a result of tighter crystal packing in the new unit cell; the β8α8 loop of chain B contacts helix α1 from an adjacent tetramer. Asp235 on the β8α8 loop, which corresponds to the metal ligand in metal-dependent KDO8PSs, is pointing away from the active site and is hydrogen bonding to His228. Although the orientation observed here is an unproductive conformation and is likely present only because of the absence of substrates and because of crystal packing effects, it almost certainly represents an orientation sampled in solution in the absence of substrates. The tighter packing between tetramers has also ordered the usually conformationally variable β2α2 loop into a common conformation in all four monomers, with

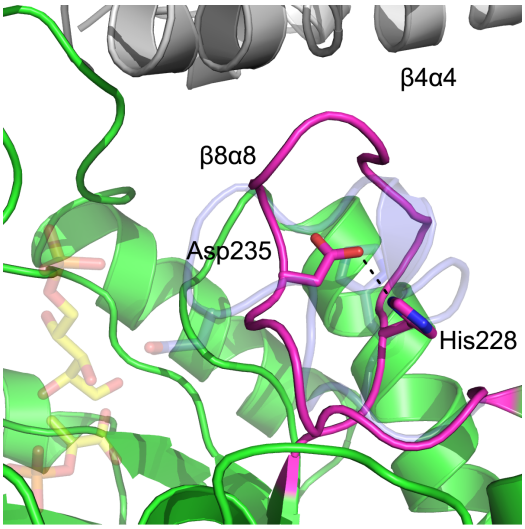


Figure 5.6: The structure of *Nme*KPRS/L7trun (green) showing the β8α8 loop (magenta). The β8α8 loop (purple), PEP and A5P (both yellow) are from the superimposed structure of *Aae*KDO8PS (PDB code 2NX3). The adjacent tetramer is coloured white.

Table 5.3: Crystal parameters, data collection, and refinement statistics for *Nme*Q202A and *Nme*S211A.

	*Nme*Q202A	*Nme*S211A
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.84, 85.74, 163.30	82.19, 85.83, 163.30
Resolution range (Å)	30.56–2.05 (2.12–2.05)	36.87-âĂŞ1.90 (1.96–1.90)
No. of measurements	305089	294670
No. of unique reflections	72315	86379
Redundancy	4.22 (4.17)	3.41 (3.25)
Completeness (%)	99.3 (99.9)	93.7 (94.3)
$I/\sigma(I)$	4.8 (1.6)	5.2 (1.6)
R_{merge}	0.067 (0.383)	0.062 (0.376)
Wilson B value (Å^2)	34.8	32.8
	Refinement	
Resolution (Å)	30.60–2.05	36.87–1.90
R_{cryst}	0.2049	0.2203
R_{free}	0.2474	0.2584
Chain length	280	280
Observed number of residues	250 + 249 + 253 + 254 residues; 7739 atom sites	250 + 251 + 252 + 251 residues; 7733 atom sites
No. of water molecules	534	586
No. of others	1 Cl^-, 1 Na^+	1 Cl^-, 1 glycerol, 1 Na^+
Mean B (Å^2)		
Protein	39.12	41.72
Water	39.52	38.43
Other	39.51	44.08
RMSD from target values		
Bond lengths (Å)	0.010	0.012
Bond angles (°)	1.102	1.215
Dihedral angles (°)	5.376	5.274
Ramachandran		
Most favoured (%)	93.3	93.9
Allowed (%)	5.5	5.1
Generously allowed (%)	0.9	0.7
Disallowed (%)	0.2	0.3
PDB code	3STE	3STF

Table 5.4: Crystal parameters, data collection, and refinement statistics for *Nme*L7trun and *Nme*KPRS/L7trun.

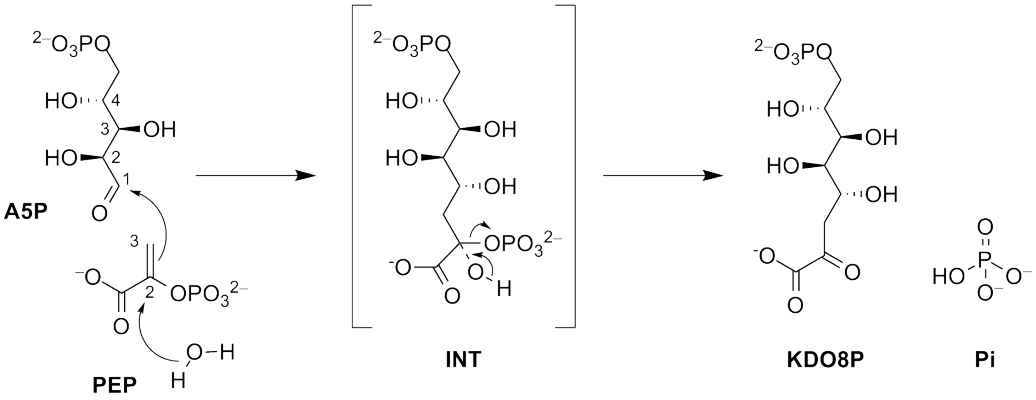
	*Nme*L7trun	*Nme*KPRS/L7trun
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.78, 85.76, 163.22	81.69, 104.05, 149.62
Resolution range (Å)	37.99–1.91 (2.01–1.91)	38.01–2.20 (2.32–2.20)
No. of measurements	522701	480225
No. of unique reflections	88598	65469
Redundancy	5.9 (4.9)	7.3 (7.4)
Completeness (%)	99.5 (99.1)	100 (100)
$I/\sigma(I)$	8.2 (1.9)	7.4 (1.8)
R_{merge}	0.062 (0.313)	0.087 (0.347)
Wilson B value (Å^2)	22.2	24.0
	Refinement	
Resolution (Å)	36.56–1.91	35.99–2.20
R_{cryst}	0.2092	0.2083
R_{free}	0.2453	0.2448
Chain length	269	268
Observed number of residues	249 + 254 + 260 + 259 residues; 7906 atom sites	260 + 268 + 260 + 258 residues; 8138 atom sites
No. of water molecules	699	574
No. of others	7 (1×0.67) Cl^-, 1 Na^+	12 Cl^-
Mean B (Å^2)		
Protein	31.07	24.86
Water	33.13	24.08
Other	33.52	28.50
RMSD from target values		
Bond lengths (Å)	0.011	0.013
Bond angles (°)	1.158	1.267
Dihedral angles (°)	5.107	5.427
Ramachandran		
Most favoured (%)	94.2	93.3
Allowed (%)	4.9	5.8
Generously allowed (%)	0.6	0.7
Disallowed (%)	0.3	0.2
PDB code	3STC	3STG

the conserved Arg59 (Arg60 in wild-type enzyme and mutants retaining the KANRS motif) always pointing toward the active site. In contrast, in the structure of *Nme*L7trun, in three of four chains, Arg60 is in an unproductive orientation, hydrogen bonding to Asp120 of the adjacent monomer, which normally is observed to hydrogen bond to Arg69 at the end of the β2α2 loop.

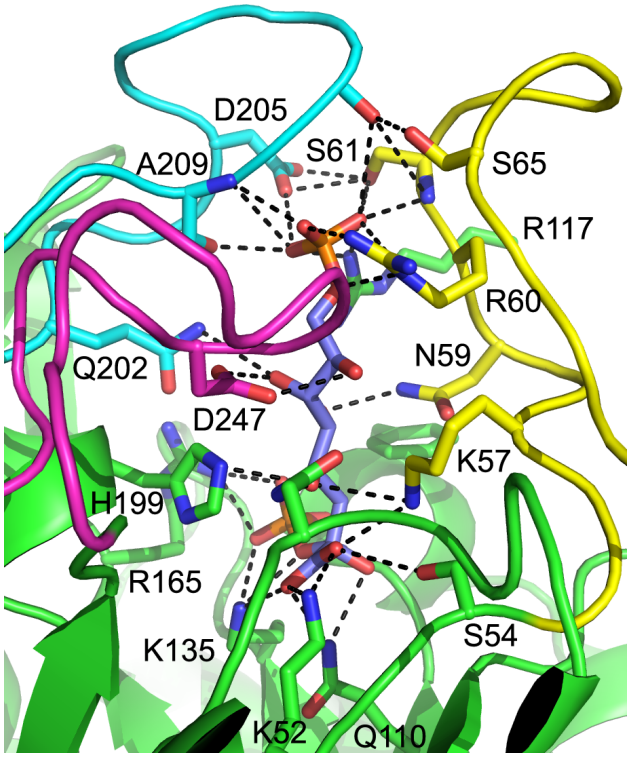
5.5 Modelling of the reaction intermediate

To examine the possible interactions between the substrates and the residues of the β7α7 loop, the proposed tetrahedral intermediate for the KDO8PS reaction was modelled into *Nme*KDO8PS using induced-fit docking, by Parker research group member Dr Wanting Jiao. The *R* and *S* isomers of the tetrahedral intermediate were modelled into the active site of *Nme*KDO8PS, and both the poses with the correct PEP orientation were from the *S* isomer of the tetrahedral intermediate. In the best pose, the PEP part of the molecule interacts with residues Arg165, Lys135, Gln110, Lys52, and Lys57 and the backbone of Ala113 (Figure 5.7). The A5P phosphate group interacts with Arg60 and Ser61 (from the β2α2 loop), Asp205, Ser211, and the backbone of Ala209 (from the β7α7 loop), and Arg117 (extending from the adjacent subunit). The hydroxyl groups derived from A5P are held in place by hydrogen bonding to Asn59, Asp247, and Gln202. The A5P aldehyde group, now converted to a hydroxyl group in the tetrahedral intermediate, interacts with Lys57, Asn23, and His199. The hydroxyl group that corresponds to the incoming water molecule forms a hydrogen bond with the side chain of Gln110. These interactions bear close correspondence to those modelled from crystallographic observations for the intermediate in *Aae*KDO8PS (PDB entry 2NX3, chain D).[24]

It is clear from the induced-fit docking that the interactions between residues arising from the extended β7α7 loop and A5P are extensive and involve more residues of this loop than the absolutely conserved Gln202 and Ser211. In addition, hydrogen bonds between the β7α7 and β2α2 loops are



(a)



(b)

Figure 5.7: (a) The KDO8PS catalysed reaction. The existence of the tetrahedral intermediate has been demonstrated by mass spectral studies.[22] (b) Proposed tetrahedral intermediate (purple) for the KDO8PS reaction modelled into the active site of *Nme*KDO8PS. The carbon atoms of the β7α7 loop are coloured cyan, those of the β8α8 loop magenta, and those of the β2α2 loop yellow.

observed. Specifically, Ser61 hydrogen bonds to Asp209, and Ser65 interacts with the main-chain carbonyl oxygen of Ala209.

5.6 Discussion

Variants of *Nme*KDO8PS and *Afe*KDO8PS, created by excising the extended β7α7 loop, were found to have severely attenuated catalytic activity. This compromise of activity was far more significant than that which resulted from the removal of the side chains of the two absolutely conserved residues, Gln and Ser, of this loop. This loop has little effect on PEP binding, but in line with the experimental studies that indicate the importance of the extended β7α7 loop for A5P binding, modelling studies predict that the β7α7 loop interacts extensively with A5P and also helps to buttress the position of the β2α2 loop. This latter loop supports multiple key interactions with A5P and provides a catalytically essential Lys residue as discussed in the previous chapter.

The β7α7 loop is conserved for all KDO8PSs regardless of their metal dependency; however, the relative importance of the β7α7 loop for efficient catalysis and binding of A5P does appear to depend on the metal dependency of the enzyme. The metal-dependent *Afe*KDO8PS was noticeably less compromised by the truncation of the β7α7 loop than the metal-independent *Nme*KDO8PS, or for that matter by mutation of the conserved residues. The metal ion appears then to play a role in supporting the correct placement of A5P for reaction, and it is conceivable the metal-dependent enzymes may be less reliant on the extended β7α7 loop for accurate A5P binding. Together, these observations appear to reflect an evolution of the metal-independent forms from a metal-dependent ancestor, which in turn evolved from the metal-dependent type Iβ DAH7PSs.

What remains striking and unexplained is the fact that no extant DAH-7PS has an extended β7α7 loop and no extant KDO8PS has a truncated β7α7 loop. The absence of this loop has little effect on PEP binding, but

its presence is associated with a change in substrate specificity, from the four-carbon E4P for DAH7PS to the five-carbon A5P for KDO8PS. The longer and more flexible A5P substrate may require additional contacts to competently bind to the enzyme and to help ensure correct substrate selection. DAH7PS, on the other hand, uses a less flexible β2α2 loop (compared to its counterpart in KDO8PS) to provide the necessary contacts to ensure effective E4P binding. The presence of the extended β7α7 loop in KDO8PS might indicate that the determining factor in the evolution of KDO8PS was the acquisition of the 11-residue β7α7 loop insertion. Curiously, the KDO8PS from *H. pylori* has a second 12-residue extension in the β6α6 loop, indicating, perhaps, that loop extension is genetically facile.

In crystal structures of both metal-dependent and metal-independent KDO8PSs, the β7α7 loop is often disordered and becomes visible only in structures with both PEP and A5P bound, as in the KDO8PS from the thermophile *A. aeolicus*.[72] The apparent mobility of the loop maintains easy access to the active site for the substrates, as well as exit for the products, and when ordered and closed helps arrange and select the substrate. The extended loop is an additional tool for helping to orient the substrate properly, and the interactions it provides also act in concert with other active-site contacts, particularly the essential ones provided by the β2α2 loop. The entropic cost of ordering a mobile β7α7 loop as part of the catalytic cycle probably represents a compromise between facile active-site access and the need for added strategies for binding A5P, as is apparent in the Michaelis constants for metal-independent *Nme*KDO8PS compared to those of metal-dependent *Afe*KDO8PS, to ensure efficient catalysis and accurate substrate selection.

The distinctive ordering of the otherwise disordered β8α8 loop observed in one subunit in the structure of KPRS/L7trun *Nme*KDO8PS provides, perhaps, a clue into controlling the entropic penalty of the ordering of long apparently flexible loops on substrate binding. The β8α8 loop of this subunit traces approximately the path observed in substrate-bound *Aae*KDO8PS,[72] but changes in side-chain conformations of absolutely conserved residues Asp235 and His228 on loop β8α8 and also of Glu26 (*Nme*KDO8PS numbering)

lead to alternative contacts with one another, compared to those contacts observed in substrate-bound *Aae*KDO8PS. Although the orientation observed here for Asp235 is an unproductive conformation and is likely present only because of the absence of substrates and is locked-in because of crystal-packing effects, it almost certainly represents an orientation sampled in solution in the absence of substrates. The different uses of absolutely conserved residues by absolutely conserved residues on loop β8α8 of *Nme*KDO8PS and *Aae*KDO8PS provide a plausible resolution of the paradox of the huge conformational entropy penalty when a flexible loop potentially sampling a multitude of conformations becomes locked into one productive conformation upon substrate binding: should this long β8α8 loop sample only a small number of defined but very different conformations, as contrasted with a more continuous multitude of conformations, then the entropic penalty on binding substrates is greatly reduced. This reduced penalty need not increase if the conformations of one long potentially flexible loop are then correlated with another, such as loops β2α2 and β7α7 in non-truncated forms. It should be noted that more than a couple of conformations, even if very well ordered, are generally difficult to discern by X-ray diffraction techniques.

Interestingly, in Iβ DAH7PS, where the β7α7 loop is short and not extended, a conserved Arg on the β6α6 loop partially occupies a similar space to that of the extended β7α7 loop in KDO8PSs, specifically that of the conserved Gln (Figure 5.8). In the structures of the Iβ DAH7PSs *Pfu*DAH7PS, *A. pernix* DAH7PS (*Ape*DAH7PS) and *Tma*DAH7PS this Arg interacts with a conserved (in Iβ DAH7PS) Gln on the β8α8 loop, three residues downstream of the metal-ligand Asp. It would seem that in Iβ DAH7PS the Arg-Gln interaction helps to secure the position of the β8α8 loop relative to that of the β6α6 loop. In KDO8PS, where the extended (and flexible) β7α7 loop occupies a similar position, there are no equivalent interactions between it and the β8α8 loop. Intriguingly, in both type Iα and II DAH7PSs this Arg and Gln are not present and nor do there appear to be any compensatory differences.

In summary, while the extended β7α7 loop is not absolutely essential for catalysis by metal-dependent and metal-independent KDO8PS enzymes,

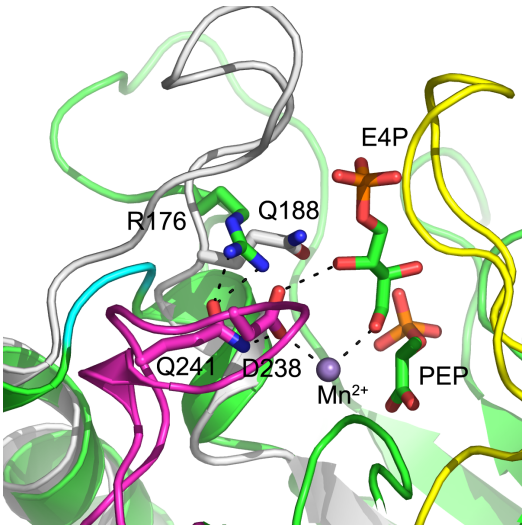


Figure 5.8: Structure of *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]) showing the conserved Arg residue (R176) on the β6α6 loop which hydrogen bonds a conserved Gln on the β8α8 loop. The β7α7 loop of *Aae*KDO8PS is superimposed (PDB code 2NX3, coloured white) with the conserved Gln (Q188) shown which occupies a similar position to R176 in *Pfu*DAH7PS. The β2α2 loop is coloured yellow, and the β7α7 loop cyan, and the β8α8 loop magenta.

its presence is necessary for efficient catalysis. This requirement for catalytic efficiency is not solely due to the interactions that involve the side chains of the absolutely conserved Ser and Gln residues of this loop, but rather appears to demand the presence of the loop as a whole.

Chapter 6

The role of a conserved Asp on the *re* face of PEP

6.1 Introduction

One subtle difference to the active-site architecture between DAH7PS and KDO8PS is the identity of a residue on the *re* face of PEP. In DAH7PS, the residue is conserved as a Glu and in KDO8PS the equivalent residue, in sequence and space, is conserved as an Asp (Figure 6.1). The residue is located on the β3-strand immediately before the beginning of the β3α3 loop in the PEP binding pocket, and in crystal structures of both enzymes where PEP is bound, makes no direct contact with PEP (Figure 6.2).

In DAH7PS, the Glu has been proposed to be part of a proton relay chain that links deprotonation of a water molecule, which attacks the *re* face of PEP, with protonation of the aldehydic carbon of E4P (Figure 6.2a and 6.3c).[74] However, for KDO8PS several possibilities have been suggested for how the water molecule that attacks PEP is activated, and these depend on whether the water molecule attacks on either the *re* or *si* face of PEP.[19,24,27,72,73,78,79] What supplies the proton for the aldehydic carbon of A5P in KDO8PS is also

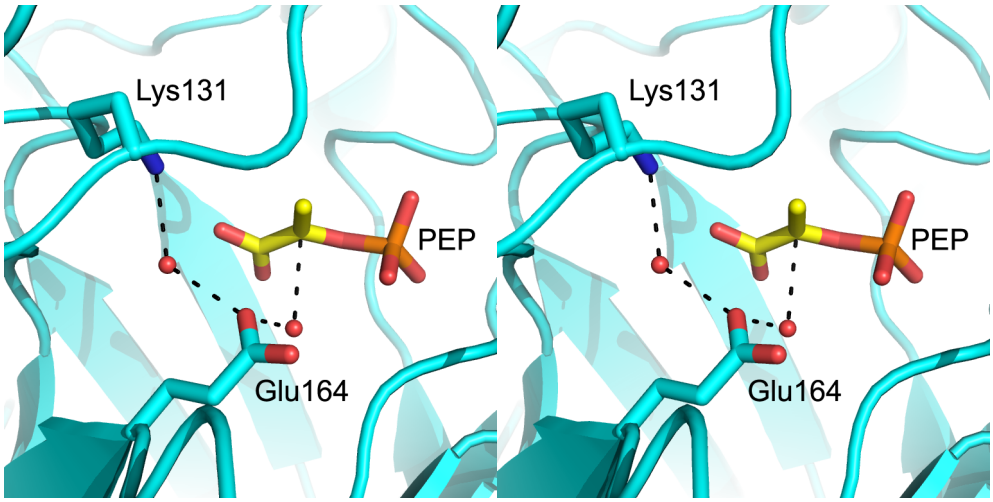
```
PfuDAH7PS    ADEYGLVTVTEVMDTRHVELV
ApeDAH7PS    GDEAGLPVVTEVLDPRHVETV
TmaDAH7PS    ADKYGMYVVTEALGEDDLPKV
SpyDAH7PS    CQEFGLLSVSEIMSERQLEEA
NmeKDO8PS    KAEFGIPVITDVHEPHQCQPV
EcoKDO8PS    KQTFGVKIITDVHEPSQAQPV
AaeKDO8PS    KEEFGLKITTDIHESWQAEPV
AfeKDO8PS    RREVGVPVVTDVHEKEDVSAV
                  *:   ::      .   .
```

Figure 6.1: Partial sequence alignment of Iβ DAH7PSs and KDO8PSs generated using ClustalX. The conserved Glu/Asp is highlighted grey.

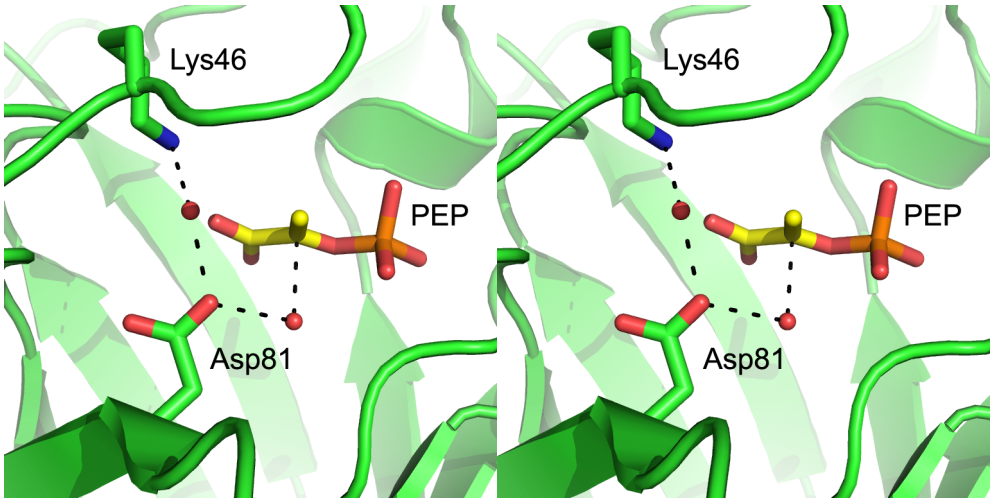
unclear. The source will depend on which, if any, of the crystallographically observed conformations of A5P (which vary in the position of the A5P aldehyde group) is catalytically competent.

If A5P is oriented with its aldehyde closest to the PEP phosphate (as has been observed in crystal structures[72]), and the phosphate group of PEP is in a monoanionic rather than dianionic state, then as the reaction proceeds the phosphate group could supply the proton to the nascent negative charge on the aldehyde of A5P (Figure 6.3a and 6.3b).[69] It was suggested that this charge state of the PEP phosphate could be favoured in KDO8PS because of a nearby conserved hydrophobic Phe residue. In DAH7PS, a conserved Arg, which forms a salt bridge with the phosphate group of PEP, occupies the equivalent space. If the PEP phosphate group is the proton source, unlike for DAH7PS, in KDO8PS the conserved Asp would have no direct role to play in the catalysed reaction, and only a potentially limited influence on which face of PEP the water molecule attacked. Additionally, the reason for the conservation of the Asp residue, and the conserved difference in residue choice between KDO8PS and DAH7PS would remain an enigma.

An alternative scenario, and that which is similar to DAH7PS, is A5P is oriented with its aldehyde closest to the metal-ion binding site (or remnants of) and toward the PEP carboxylate rather than PEP phosphate (Figure 6.2b). As for the alternative binding arrangement, water could attack PEP on either face, but this arrangement makes possible for the same proton



(a)



(b)

Figure 6.2: (a) Stereoview of the structure of *Tma*DAH7PS (PDB code 1RZM) showing the conserved Glu (Glu164) and the hydrogen-bond network connecting the nucleophilic water to the proton-donor Lys (Lys131). (b) Stereoview of the structure of *Aae*KDO8PS (PDB code 2NX3) showing the conserved Asp (Asp81) and the equivalent hydrogen-bond network. The carbon atoms of PEP are coloured yellow.

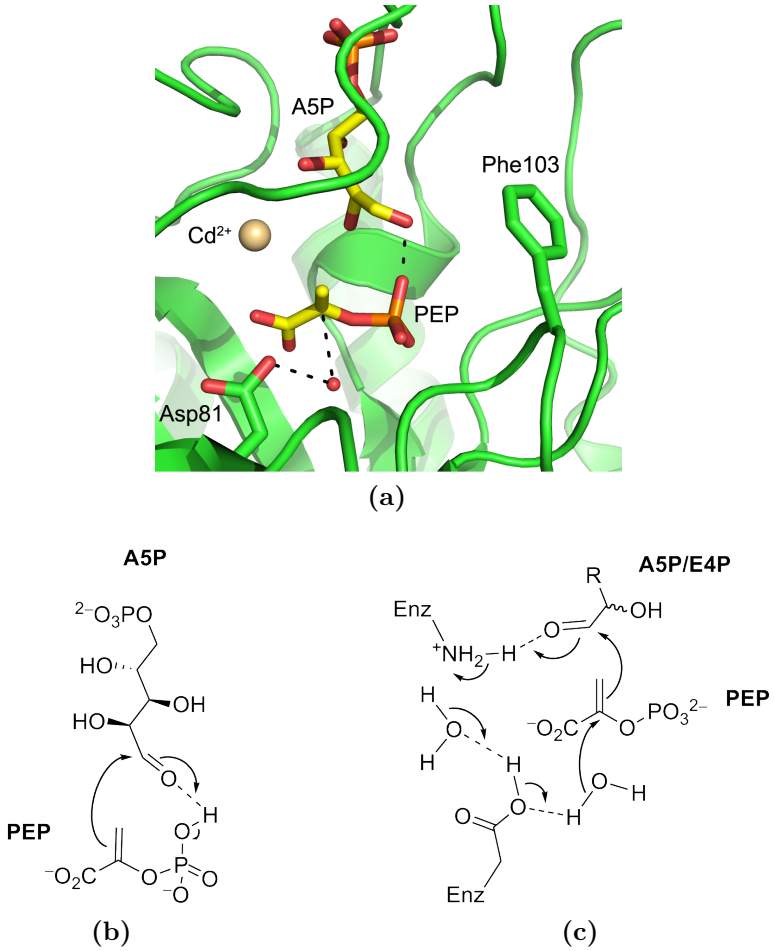


Figure 6.3: The alternative proton sources for the aldehyde of A5P/E4P. (a) The structure of *Aae*KDO8PS (PDB code 1FWW) showing A5P in a conformation that maintains the *re* face of A5P toward the *si* face of PEP but in which the aldehyde is within hydrogen-bonding distance of the PEP phosphate. (b) Schematic representation of a monoanionic PEP-phosphate providing a proton to the aldehyde of A5P. (c) Possible proton relay chain for both KDO8PS and DAH7PS that links protonation of the aldehyde with deprotonation of the nucleophilic water.

relay chain that is thought to occur in DAH7PS to also occur in KDO8PS (Figure 6.3c). If this chain does operate in KDO8PS, then the water must attack the *re* face of PEP. The same proton relay chain operating in both enzymes is consistent with the Lys residue at the terminus of the proton relay being essential for catalysis in both KDO8PS (as found in Chapter 4) and DAH7PS.[97] The only difference in the relays between the two enzymes would be the involvement of the Asp in KDO8PS rather than the Glu in DAH7PS. This would appear to be a negligible difference, as the side chain of Asp, while one-carbon shorter than Glu, presents the same carboxylic acid functional group. However, a subtle change like this suggests, perhaps, that the adaptation to use Asp may be the result of co-evolution, and coupled with other residue differences between the two enzymes.

A third possibility, counter to both DAH7PS and KDO8PS sharing the same proton relay chain mechanism, is the potential for the metal-ligand Asp to be involved with both activating the nucleophilic water (on the *si* face of PEP), but also supplying a proton to the aldehyde group of A5P.[29] If this is the case, then again the role and reason for the conservation of Asp in KDO8PS seems unclear.

In this chapter the role of the conserved Asp in KDO8PS is investigated.

6.2 Preparation of Asp mutants

To study the role of the conserved Asp residue, enzymes were created where it was mutated to Ala, Asn and Glu, in both the metal-dependent *Afe*KDO8PS (creating *Afe*D90A, *Afe*D90N and *Afe*D90E) and the metal-independent *Nme*-KDO8PS (creating *Nme*D92A, *Nme*D92N and *Nme*D92E). The Asp-to-Ala mutants were created to remove the carboxylic acid functionality and the Asp-to-Asn mutants were created to assess the effect of the isosteric amide functionality on enzyme function. The Asp-to-Glu mutations were created to change the residue to that found in DAH7PS, and hence to probe the importance of side-chain length versus functionality.

The mutant proteins of *Afe*KDO8PS and *Nme*KDO8PS were purified using the respective wild-type enzyme purification procedures.[39,84] CD was used to assess the structural integrity of the *Afe*KDO8PS mutant proteins, the spectra of which were identical to each other and wild-type protein (Figure 6.4). The melting temperatures were measured by DSF for the mutant proteins of both *Afe*KDO8PS and *Nme*KDO8PS in the presence of different additives (Table 6.1). Generally, the *Afe*KDO8PS mutants were more stable than wild type by 2 ℃ to 5 ℃, but the presence of Cd^{2+} caused a smaller increase in T_m than for wild type. In contrast, the *Nme*KDO8PS mutants were all slightly less stable than wild type (by 1 ℃ to 3 ℃) in all conditions.

6.3 Kinetic characterisation

The catalytic activity of the six mutant enzymes was assessed (Table 6.2). While there was no measurable activity for either *Afe*D90A or *Afe*D90N,

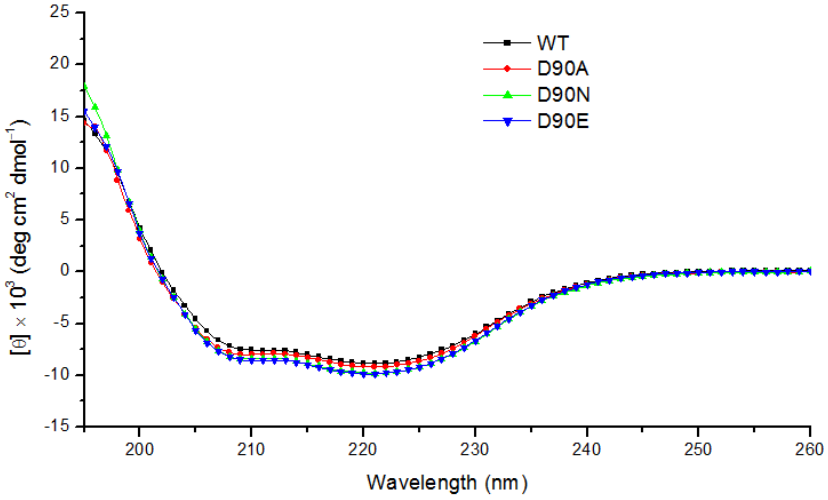


Figure 6.4: CD spectrophotometry of *Afe*KDO8PS wild-type and mutants. Wild-type: black squares; *Afe*D90A: red circles; *Afe*D90N: green triangles; *Afe*D90E: blue triangles. Every second datum point is plotted.

Table 6.1: The effect of additives on T_m for wild-type and Asp mutants of *Afe*KDO8PS and *Nme*KDO8PS.

KDO8PS	Additives	T_m (°C)	Diff.	Diff. to WT
*Afe*D90A	No additive	70 ± 1		2
	Mn^{2+}	72 ± 1	2	4
	Cd^{2+}	76 ± 1	6	−6
	PEP	73 ± 1	3	4
*Afe*D90N	No additive	70 ± 1		2
	Mn^{2+}	72 ± 1	2	4
	Cd^{2+}	76 ± 1	6	−6
	PEP	74 ± 1	4	5
*Afe*D90E	No additive	72 ± 1		4
	Mn^{2+}	73 ± 1	1	5
	Cd^{2+}	78 ± 1	6	−4
	PEP	73 ± 1	1	4
*Nme*D92A	No additive	55 ± 1		−3
	Mn^{2+}	56 ± 1	1	−3
	Cd^{2+}	42 ± 1	−13	0
	PEP	60 ± 1	5	0
*Nme*D92N	No additive	57 ± 2		−1
	Mn^{2+}	58 ± 2	1	−1
	Cd^{2+}	43 ± 3	−14	1
	PEP	59 ± 1	2	−1
*Nme*D92E	No additive	57 ± 1		−1
	Mn^{2+}	56 ± 1	−1	−2
	Cd^{2+}	44 ± 2	−13	2
	PEP	57 ± 1	0	−3

*Afe*D90E was active. For *Afe*D90E the values of K_m^{PEP} and K_m^{A5P} were both significantly higher (≈45 and ≈80 times respectively) than wild-type *Afe*-KDO8PS, and k_{cat} was reduced to only 25 percent of the wild-type value.

In contrast to *Afe*KDO8PS, all three D92 mutants of *Nme*KDO8PS were active. Compared to *Afe*D90E, the increase in K_m^{PEP} for *Nme*D92E was more modest (33 times), though the effect on K_m^{A5P} was more severe (90 times larger). Unlike for *Afe*D90E which remained relatively active, the k_{cat} for *Nme*D92E was a mere twelfth of the wild-type value.

Table 6.2: Kinetic parameters for wild-type and Asp mutants of *Nme*KDO8PS and *Afe*KDO8PS.

KDO8PS	K_m^{PEP} (µM)	K_m^{A5P} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{PEP} ($s^{-1}\,mM^{-1}$)	k_{cat}/K_m^{A5P} ($s^{-1}\,mM^{-1}$)
*Nme*WT	2.5 ± 0.2	12.0 ± 0.5	8.0 ± 0.1	3200 ± 300	660 ± 40
*Nme*D92A[a]	41 ± 3		>0.14		
*Nme*D92N	10 ± 1	23 ± 2	0.0244 ± 0.0006	2.4 ± 0.3	1.1 ± 0.1
*Nme*D92E	83 ± 14	1101 ± 126	0.67 ± 0.04	8 ± 2	0.6 ± 0.1
*Afe*WT	12 ± 1	22 ± 2	4.8 ± 0.1	400 ± 40	220 ± 20
*Afe*D90A	–	–	–		
*Afe*D90N	–	–	–		
*Afe*D90E	544 ± 20	1774 ± 251	1.14 ± 0.02	2.1 ± 0.1	0.6 ± 0.1

Enzymes for which no activity was measurable ($k_{cat} < 0.001\ s^{-1}$) are denoted with dashes.
[a] In experiments to determine K_m^{A5P} V_{max} was not obtained and hence a value for K_m^{A5P} could not be calculated. The value for K_m^{PEP} was measured at a sub-saturating concentration of A5P. k_{cat} calculated from the maximum observed rate.

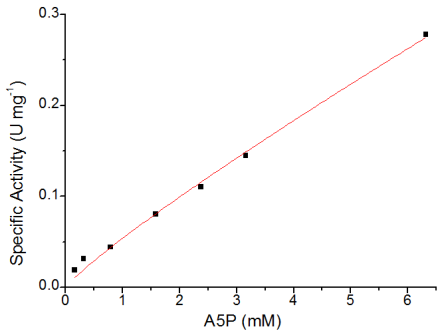


Figure 6.5: Saturation curve for *Nme*D92A with increasing A5P concentration.

The Michaelis constants for both PEP and A5P were much less affected for the *Nme*D92N enzyme, however enzyme activity was severely attenuated. For *Nme*D92A, K_m^{PEP} was increased relative to that of wild type, but K_m^{A5P} was unable to be determined; there was no reduction in the rate increase with concentrations of A5P up to 6 mM (Figure 6.5).

6.4 Measurement of PEP binding

ITC was used to measure the binding interactions of PEP to the mutant enzymes. Although *Afe*D90A and *Afe*D90N were both kinetically inactive, PEP still bound to each enzyme. The PEP dissociation constant for each protein was only modestly increased relative to wild-type *Afe*KDO8PS ($K_{\mathrm{d}}^{\mathrm{D90A}}$ is $9 \pm 4\,\mu\mathrm{M}$, $K_{\mathrm{d}}^{\mathrm{D90N}}$ is $13.7 \pm 0.3\,\mu\mathrm{M}$ and $K_{\mathrm{d}}^{\mathrm{WT}}$ is $5.0 \pm 0.1\,\mu\mathrm{M}$) (Figure 6.6a and 6.6b). The PEP binding of *Nme*D92A was measured to have a similar PEP dissociation constant to both wild-type *Nme*KDO8PS, *Afe*D90A and *Afe*D90N ($K_{\mathrm{m}}^{\mathrm{D92A}}$ of $13 \pm 2\,\mu\mathrm{M}$) (Figure 6.6c).

Attempts were made to measure PEP binding for *Afe*D90E, however no enthalpy of binding was observed, and the protein had precipitated at the completion of all experiments. Similarly, for both *Nme*D92N and *Nme*D92E, no enthalpy of binding was observed. However, the three enzymes are all kinetically active, which allowed the value of $K_{\mathrm{m}}^{\mathrm{PEP}}$ to be measured for each. It is unclear whether the precipitation of the proteins caused a lack of enthalpic binding or not.

Variable stability was experienced during the ITC experiments for all six Asp mutant proteins. This was confounding and made it difficult to repeat and improve experiments as sometimes the protein would precipitate and other times it would not. The instability of these mutant proteins does not correlate with the melting temperatures measured by DSF, for which no gross changes in T_{m} values were measured.

6.5 Crystallography

Diffraction data were collected from crystals of *Nme*D92A, *Nme*D92N and *Nme*D92E (Table 6.3 and 6.4). As for the wild-type protein, the asymmetric unit for all three structures contains one complete tetramer. In the structure of *Nme*D92A the only notable change in the final model is the truncation of the Asp side chain to Ala (RMSD for superposition of Cα atoms on wild-type

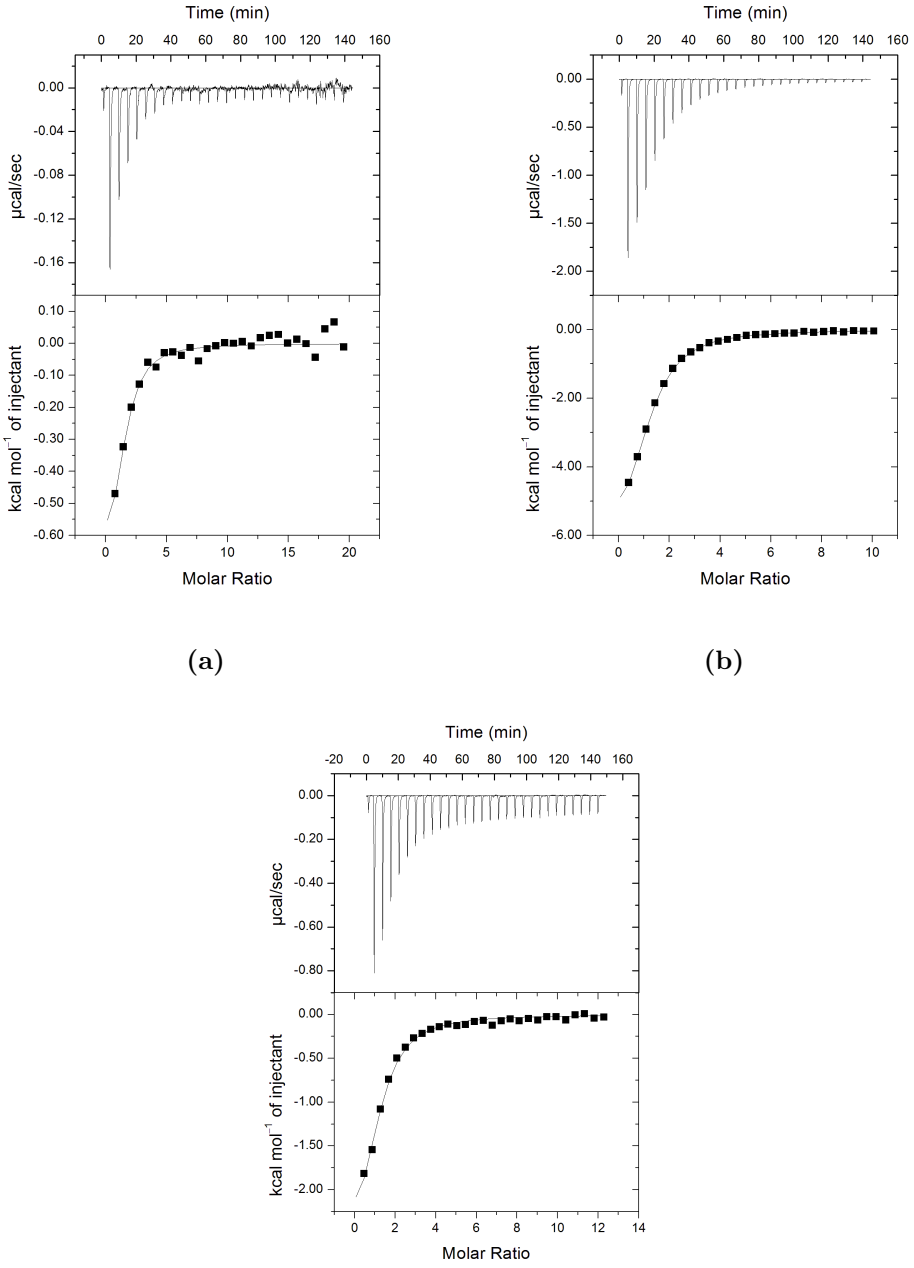


Figure 6.6: The interaction between (a) *Afe*D90A, (b) *Afe*D90N and (c) *Nme*D92A and PEP, quantified by ITC.

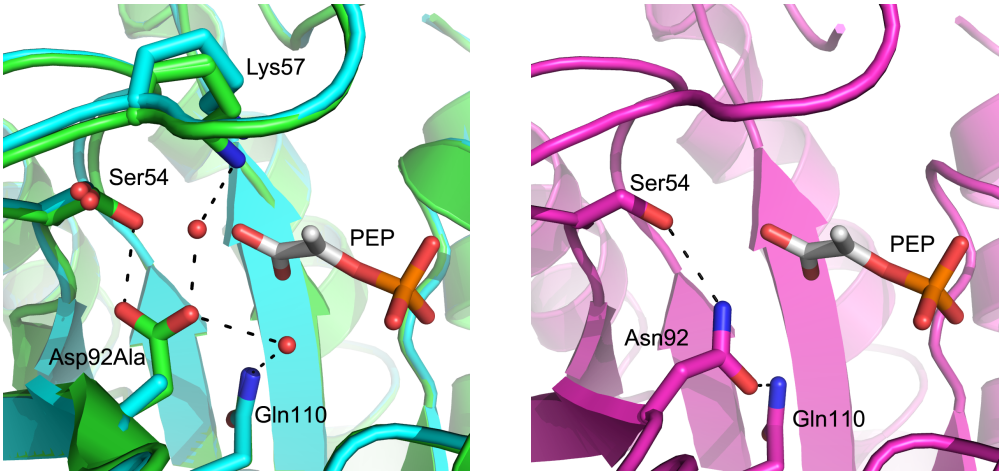


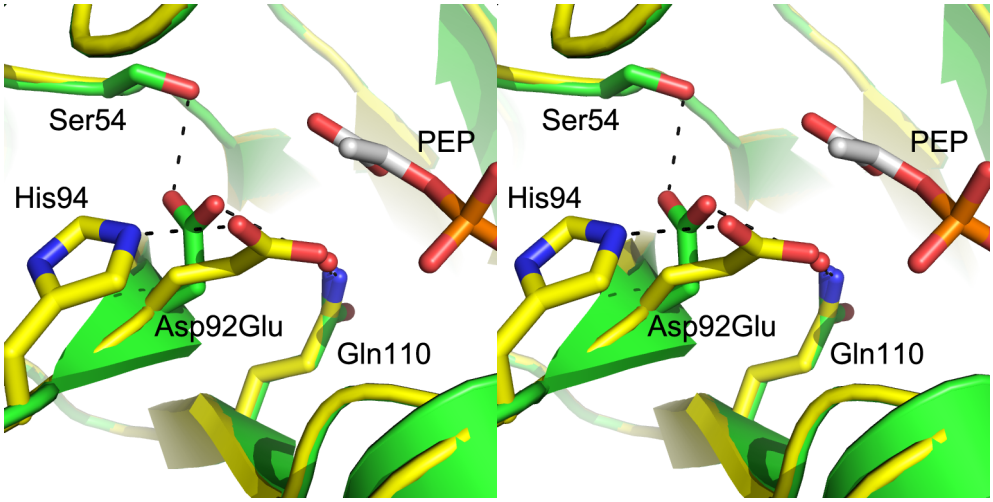
Figure 6.7: (a) The structure of *Nme*D92A (coloured cyan) superimposed onto the structure of wild-type *Nme*KDO8PS (coloured green). The water molecules shown are from the wild-type model. (b) The structure of *Nme*D92N. In both sub-figures PEP (carbon atoms coloured white) is superimposed from the structure of *Aae*KDO8PS (PDB code 1FWW).

*Nme*KDO8PS is 0.250 Å) (Figure 6.7a). No water molecules were observed in the space formerly occupied by the Asp side chain.

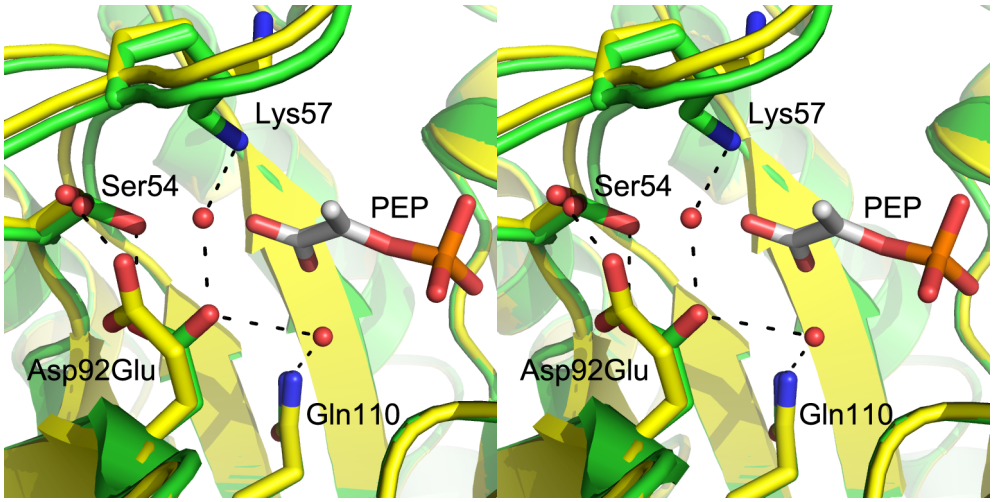
The structure of *Nme*D92N is also little different to the wild-type model other than for the introduced mutation (RMSD Cα on wild type is 0.222 Å). The $N^{\delta 2}$ and $O^{\delta 1}$ of the D92N side chain hydrogen-bond, respectively, to Ser54 and Gln110, which are in unchanged positions compared to the wild-type model (Figure 6.7b). In wild-type *Nme*KDO8PS, D92 also hydrogen bonds to Ser54 whereas the nucleophilic water hydrogen bonds to Gln110. In the structure of *Nme*D92N the *re*-face water is not observed.

The structure of *Nme*D92E was very similar in structure to wild-type *Nme*KDO8PS except for the Asp-to-Glu mutation (RMSD Cα on wild type is 0.303 Å). However, variation in the D92E side-chain conformation was observed in different protein chains. In chains A and D, the side chain of the Glu terminates where the *re*-face water in the wild-type structure is located (Figure 6.8a). In these chains, D92E also hydrogen bonds to His94 (which is

in an unchanged position relative to the wild-type structure, and conserved in KDO8PS). In chains B and C of the model, the side-chain conformation is bent, presenting the carboxylic acid functional group in a similar position to that of the wild-type Asp residue, and hence D92E hydrogen bonds with Ser54 (Figure 6.8b). In neither alternative conformation would the introduced Glu seem to be available to partake in a proton relay chain.



(a)



(b)

Figure 6.8: Stereoviews of the structure of *Nme*D92E (coloured yellow). (a) Chain A, showing the conformation of D92E in chains A and D with hydrogen bonding to His94. (b) The conformation of D92E in chain C (and B). The structures of *Nme*D92E in both sub-figures are shown superimposed onto the structure of wild-type *Nme*KDO8PS (coloured green). The water molecules shown are from the structure of wild-type *Nme*KDO8PS. PEP (carbon atoms coloured white) is from the superimposed structure of *Aae*KDO8PS (PDB code 1FWW).

Table 6.3: Crystal parameters, data collection, and refinement statistics for *Nme*D92A and *Nme*D92N.

	*Nme*D92A	*Nme*D92N
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.87, 85.70, 163.40	82.13, 85.60, 163.19
Resolution range (Å)	42.83–1.75 (1.84–1.75)	45.36–1.85 (1.95–1.85)
No. of measurements	837037	669764
No. of unique reflections	116415	98818
Redundancy	7.2 (7.3)	6.8 (6.6)
Completeness (%)	100 (100)	100 (99.9)
$I/\sigma(I)$	4.9 (1.8)	5.3 (1.9)
R_{merge}	0.086 (0.420)	0.083 (0.393)
Wilson B value (Å^2)	21.5	23.67
	Refinement	
Resolution (Å)	40.93–1.75	41.42–1.85
R_{cryst}	0.2099	0.1948
R_{free}	0.2309	0.2202
Amino acids (chain length of 280 residues)	251 + 253 + 253 + 255 residues; 7814 atom sites	250 + 251 + 253 + 254 residues; 7809 atom sites
No. of water molecules	416	333
No. of others	1 Cl^-, 1 Na^+, 1 glycerol	1 Na^+, 1 glycerol
Mean B (Å^2)		
Protein	27.41	29.20
Water	26.54	27.87
Other	22.50	21.43
RMSD from target values		
Bond lengths (Å)	0.009	0.008
Bond angles (°)	1.278	1.240
Dihedral angles (°)	5.350	5.382
Ramachandran		
Most favoured (%)	93.1	93.6
Allowed (%)	5.8	5.5
Generously allowed (%)	0.9	0.7
Disallowed (%)	0.1	0.2

Table 6.4: Crystal parameters, data collection, and refinement statistics for *Nme*D92E.

	*Nme*D92E
Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.82, 85.96, 163.79
Resolution range (Å)	48.00–1.95 (2.06–1.95)
No. of measurements	600824
No. of unique reflections	84910
Redundancy	7.1 (7.3)
Completeness (%)	100 (100)
$I/\sigma(I)$	4.0 (2.1)
R_{merge}	0.112 (0.366)
Wilson B value (Å^2)	23.8
Refinement	
Resolution (Å)	46.08–1.95
R_{cryst}	0.2255
R_{free}	0.2567
Amino acids (chain length of 280 residues)	250 + 252 + 253 + 254 residues; 7807 atom sites
No. of water molecules	271
No. of others	–
Mean B (Å^2)	
Protein	28.62
Water	23.88
Other	–
RMSD from target values	
Bond lengths (Å)	0.010
Bond angles (°)	1.336
Dihedral angles (°)	5.422
Ramachandran	
Most favoured (%)	92.3
Allowed (%)	6.7
Generously allowed (%)	0.8
Disallowed (%)	0.2

6.6 Discussion

The function of the KDO8PS-conserved Asp was probed by mutation to Ala, Asn and Glu in *Nme*KDO8PS and *Afe*KDO8PS. The mutant proteins were expressed and purified and were identical in structure to the respective wild-type proteins. The mutations were either completely deleterious to enzyme activity (in the case of D90A and D90N of the metal-dependent *Afe*KDO8PS) or reduced the activity relative to the wild-type proteins. Although the Asp residue is located within the PEP-binding pocket, surprisingly $K_{\mathrm{m}}^{\mathrm{A5P}}$ was dramatically increased compared to the wild-type proteins for *Nme*D92E, *Afe*D90E, and *Nme*D92A.

The lack of activity for *Afe*D90A and *Afe*D90N is intriguing, given the equivalent mutations in *Nme*KDO8PS produced active enzymes, albeit with modified kinetic parameters. ITC experiments demonstrated that both enzymes were still able to bind PEP with similar dissociation constant values to that of wild-type *Afe*KDO8PS. This suggests that the lack of activity is not a result of an inability to bind PEP.

For *Nme*D92A, *Nme*D92E and *Afe*D90E, the mutation of Asp to Ala or Glu caused a large increase in the value of $K_{\mathrm{m}}^{\mathrm{A5P}}$. In the case of *Nme*D92A, the $K_{\mathrm{m}}^{\mathrm{A5P}}$ was unable to be determined, with the rate at which enzyme activity increased with A5P concentration still not declining at A5P concentrations of up to 6 mM. For both *Nme*D92E and *Afe*D90E the increase in $K_{\mathrm{m}}^{\mathrm{A5P}}$ while still appreciable (80–90 times greater), was by comparison more modest. The crystal structures of *Nme*D92A and *Nme*D92E do not suggest any changes in structure apart from those of the mutations themselves. The negative effect on A5P binding is difficult to understand as in structures of KDO8PS with PEP and A5P bound, the Asp residue is observed to have neither any direct or indirect contacts with A5P that could be important for its binding.

The relatively small increase in the Michaelis constant values for *Nme*-D92N, for both PEP and A5P (in contrast to the increases observed for the Asp-to-Glu mutant proteins), suggests that a residue of similar size to Asp

can be much more easily accommodated by *Nme*KDO8PS. The value of k_{cat} for *Nme*D92N was much smaller than the wild-type value, consistent with the native Asp residue having some role to play in the enzyme mechanism. The longer side chain of Glu appears to disrupt the binding of PEP, and, given that the binding site of A5P is partially formed by PEP, this disruption is propagated to the binding of A5P. The introduction of Glu at this position into KDO8PS may be repositioning the interactions between substrates and the enzyme active site.

Analysis of structures of both KDO8PSs and DAH7PSs indicates that the position of the main chain of the β3 strand (on which the Asp/Glu resides) is the same in both enzymes, but more generally there are few differences in the active sites. In structures of both enzymes where PEP is bound, it is bound in the same position, and the residues that are conserved between both enzymes are likewise in the same position. The only difference between the active sites of the two enzymes is the position of the Asp or Glu side chain. The side-chain position of the Glu differs between *Nme*D92E and that which is adopted in structures of Iβ DAH7PS (Figure 6.9). The Iβ DAH7PS

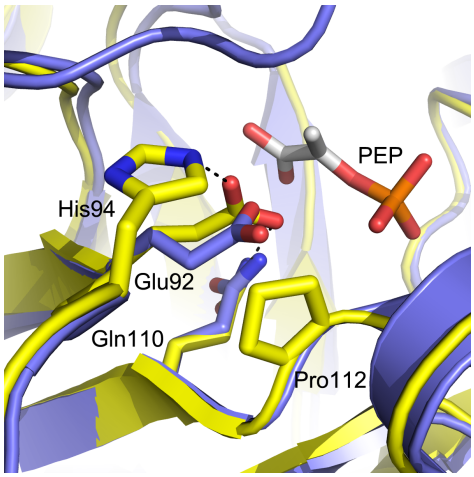


Figure 6.9: Superposition of the structure of *Nme*D92E (chain A) (coloured yellow) onto the structure of *Pfu*DAH7PS (PDB code 1ZCO, coloured blue). PEP is superimposed from the structure of *Aae*KDO8PS (PDB code 1FWW). The His at position 94 in *Nme*KDO8PS is conserved in all KDO8PSs, but the equivalent residue in Iβ DAH7PS is not conserved, and is commonly either Met, Leu or Val.

conformation of Glu is inaccessible in KDO8PS because of the presence of a conserved Pro residue at the beginning of the β4α4 loop. The same residue is conserved as Gly in DAH7PS, which allows space for the Glu to adopt the conformation observed. The Pro in KDO8PS is part of a short PAFLxR conserved sequence motif, while the Gly in DAH7PS is part of an equivalently conserved GARNxQ motif. This motif difference, along with other differences between the two enzymes are investigated in the next chapter.

6.7 Summary

The results in this chapter have identified that the conserved Asp in KDO8PS is neither important structurally or for PEP binding. The residue appears to be important for enzyme activity, with mutation (to Ala, Asn and Glu) adversely affecting the kinetic parameters for *Afe*KDO8PS and *Nme*KDO8PS. The results do not rule out the native Asp residue having a role to play, potentially as part of a proton relay mechanism, in the enzyme mechanism. However, the results also do not eliminate the possibility of a proton being supplied to the carbonyl of A5P by another method. Substitution of the Asp in KDO8PS for the DAH7PS-conserved Glu is accommodated, but could be optimised by coupling the change with mutation to other conserved differences between the two enzymes.

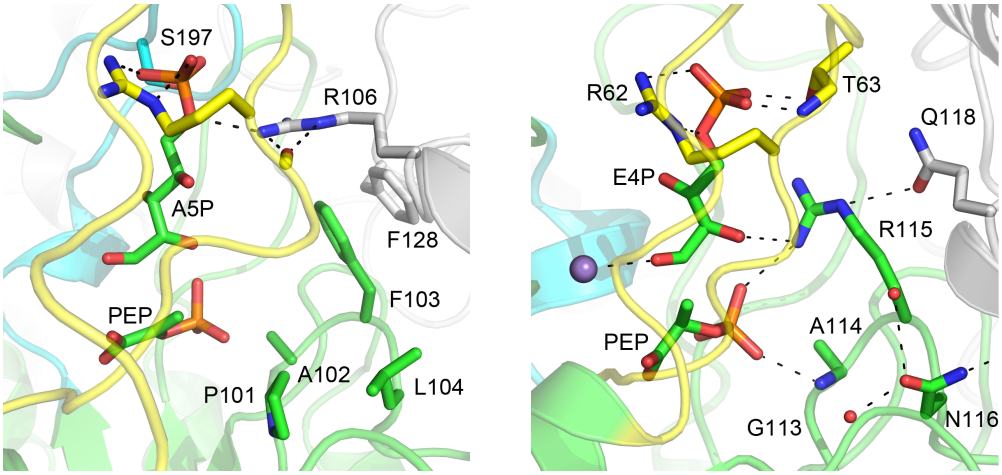
Chapter 7

Subunit interface influence on substrate selection

7.1 Introduction

The final sequence feature of KDO8PS investigated in this thesis is a short five-residue motif on the β4α4 loop, composed of the conserved residues PAFLxR. The motif is close to the area in which the PEP phosphate moiety binds in the active site and forms part of the interface with an adjacent monomer (Figure 7.1a). The Pro residue of the motif appears to mediate a bend at the top of the β4 strand at the beginning of the loop. The Phe residue interacts across the interface with a conserved Phe on the β5α5 loop from the adjacent monomer. A notable feature of the motif is the Arg residue, which interdigitates across the interface into the active site of the adjacent monomer and contributes to the binding region for the phosphate moiety of A5P.

In DAH7PS, the equivalent residues are also conserved, however in this enzyme the motif is composed of the residues GARNxQ. The Arg, which replaces the Phe, has hydrogen-bonding interactions with the phosphate



(a) **(b)**

Figure 7.1: (a) The structure of *Aae*KDO8PS (PDB code 2NX3) showing the location of the PAFLxR motif. (b) The structure of *Pfu*DAH7PS (PDB code 1ZCO, E4P modelled[59]) showing the location of the GARNxQ motif. The adjacent monomer is coloured white, the β2α2 loop yellow and the β7α7 loop cyan.

moiety of PEP and with C2-OH of E4P (Figure 7.1b). The Gln, which replaces the interdigitating Arg of KDO8PS, seems to buttress the Arg of the motif into position.

In this chapter, the role of the PAFLxR motif in KDO8PS is investigated by creating single mutations and combinations of double and triple mutations to the motif in *Nme*KDO8PS. Additionally, changes to the PAFLxR motif were combined with mutations to other areas of the active site, which have been the subjects of investigations in previous chapters, to create a protein that more closely resembles a DAH7PS than KDO8PS.

7.2 Preparation of motif mutants

Mutations to the ^{112}PAFLxR^{117} motif in *Nme*KDO8PS were created by site-directed mutagenesis (Table 7.1). Phe114 of the motif was mutated to Ala and Arg, creating *Nme*F114A and *Nme*F114R. These two constructs were created by Parker group member Dr Fiona Cochrane. The conserved Phe,

Table 7.1: The interface-area mutations made to *Nme*KDO8PS.

Construct	Mutation location
*Nme*F114A	PA**F**LxR
*Nme*F114R	PA**F**LxR
*Nme*R117A	PAFLx**R**
*Nme*R117K	PAFLx**R**
*Nme*R117Q	PAFLx**R**
*Nme*F139G	β5α5 loop
*Nme*F114R/R117A	PA**F**Lx**R**
*Nme*F114R/R117Q	PA**F**Lx**R**
*Nme*F114R/R117Q/F139G	PA**F**Lx**R** and β5α5 loop

not part of the motif, but with which Phe114 partners was mutated to Gly (creating *Nme*F139G). Gly is the conserved identity of the equivalent residue in Iβ DAH7PS. The Arg (PAFLx**R**), which interdigitates into the active site of the adjacent monomer, was mutated to Ala (creating *Nme*R117A), Gln (as found in DAH7PS) creating *Nme*R117Q, and Lys, creating *Nme*R117K. *Nme*R117A and *Nme*R117K were also created by Dr Fiona Cochrane. Variants of *Nme*KDO8PS that contained combinations of these mutations were also created: *Nme*F114R/R117A, *Nme*F114R/R117Q, and *Nme*F114R/R117Q/F139G. The latter two resemble the pairings of residues found in DAH7PS.

The mutant proteins were purified using the same procedures as those developed for the purification of wild-type *Nme*KDO8PS[39] and the melting temperatures were measured by DSF. Some of the mutated proteins were found to be more stable than wild-type *Nme*KDO8PS, whereas others were less stable (Table 7.2). *Nme*F114A was generally more stable than wild-type protein in all conditions (by 6 ℃ to 8 ℃), and interestingly, the T_m value was not affected by the presence of Cd^{2+}, which usually has a large destabilising effect for *Nme*KDO8PS. Although *Nme*F139G had similar T_m values to wild-type *Nme*KDO8PS in all conditions, like for *Nme*F114A, the T_m value was also not affected by the presence of Cd^{2+}. *Nme*F114R was generally less stable than wild-type *Nme*KDO8PS (by 3 ℃ to 5 ℃), and presence of Cd^{2+}

Table 7.2: The effect of additives on T_m for interface-area mutants of *Nme*-KDO8PS.

KDO8PS	Additives	T_m (°C)	Diff.[a]	Diff. to WT
*Nme*F114A	No additive	66 ± 1		8
	Mn^{2+}	65 ± 1	−1	6
	Co^{2+}	66 ± 1	0	8
	Cd^{2+}	64 ± 1	−2	22
	PEP	66 ± 1	0	6
*Nme*F114R	No additive	54 ± 1		−4
	Mn^{2+}	54 ± 1	0	−5
	Co^{2+}	54 ± 1	0	−3
	Cd^{2+}	43 ± 1	−11	1
	PEP	55 ± 1	1	−5
*Nme*F139G	No additive	57 ± 1		−1
	Mn^{2+}	57 ± 1	0	−2
	Co^{2+}	57 ± 1	0	0
	Cd^{2+}	54 ± 1	−3	12
	PEP	58 ± 1	1	−2
*Nme*R117Q	No additive	69 ± 1		11
	Cd^{2+}	64 ± 1	−5	12
	PEP	72 ± 3	3	12
*Nme*F114R/R117A	No additive	54 ± 1		−4
	Cd^{2+}	41 ± 1	−13	−1
	PEP	62 ± 1	8	2
*Nme*F114R/R117Q	No additive	58 ± 1		0
	Cd^{2+}	50 ± 1	−8	8
	PEP	65 ± 1	7	5
*Nme*F114R/R117Q/F139G	No additive	58 ± 1		0
	Cd^{2+}	49 ± 1	−9	7
	PEP	65 ± 1	7	5
*Nme*Quin	No additive	37 ± 1		−21
	Mn^{2+}	37 ± 1	0	−22
	Cd^{2+}	37 ± 4	0	−5
	PEP	45 ± 1	8	−15

[a] Difference with respect to no additive conditions.

remained a destabilising condition. The *Nme*R117Q protein was generally more stable than wild-type *Nme*KDO8PS (by 11 ℃ to 12 ℃), but only showed a slight destabilisation in the presence of Cd^{2+}. The melt temperatures for both double mutants (*Nme*F114R/R117A and *Nme*F114R/R117Q) and the triple mutant (*Nme*F114R/R117Q/F139G) were similar to those for wild-type *Nme*KDO8PS.

Preparation and behaviour of a putative DAH7PS

In Chapter 5 it was shown that combining the truncation of the β7α7 loop with mutation of the KDO8PS-conserved KANR(S/T) motif to KPR(S/T) as it is found in DAH7PS, in the metal-dependent *Afe*KDO8PS and metal-independent *Nme*KDO8PS, resulted in enzymes that were inactive, for both KDO8PS (use of A5P) and DAH7PS (use of E4P) activity. Herein, the *Nme*KPRS/L7trun protein was further mutated to include the PAFLxR-related mutations F114R, R117Q and F139G, in an attempt to favour the selection and use of E4P as a substrate. This construct, containing these five modifications, will be referred to as *Nme*Quin.

The three mutations were added sequentially to the *Nme*KPRS/L7trun template, first F139G, then F114R and finally R117Q. However, in attempts to add the last mutation, many difficulties were encountered. Sequencing of colonies would invariably show that the site-directed mutagenesis process had introduced unwanted errors in sequence, most commonly within the region of the primers used for mutagenesis, often in the form of deleted bases. Several strategies were used in attempts to get a construct containing the desired sequence. These included sequencing more colonies, repeating and varying the conditions of the polymerase chain reaction (PCR) reaction and using further mutagenesis to remove unwanted mutations. However, all attempts were unsuccessful, and a construct that also contained a single undesired mutation was chosen so as to not retard progress. This construct contained the undesired mutation A126T, which is located at the end of helix α4, at the opposite end of the TIM-barrel from the active site. Inspection of the

structure of *Nme*KDO8PS suggested that the mutation was unlikely to affect the properties of the protein and that preliminary studies with this protein may be worthwhile.

In initial purifications of *Nme*Quin, protein was abnormally lost during the purification and also observed to precipitate out of solution after the final SEC step. Addition of salt and PEP to all purification buffers enabled the purification to achieve similar yields to other *Nme*KDO8PS ones. The measurement by DSF of the T_{m} values in different conditions revealed that the combination of mutations had destabilised *Nme*Quin, melt temperatures were 20 ℃ less than for wild-type *Nme*KDO8PS (Table 7.2). However, the presence of PEP was particularly stabilising (more so than for wild-type *Nme*KDO8PS), explaining the positive effect of the addition of PEP to the purification buffers. It is unknown whether the instability of *Nme*Quin is caused by the combination of the (desired) mutations to the protein, and/or the presence of the undesired A126T mutation. Measurement of the CD spectrum indicated that *Nme*Quin is normally folded (Figure 7.2).

All mutant constructs characterised in this chapter eluted with similar SEC profiles, suggesting that the mutations do not affect the quaternary

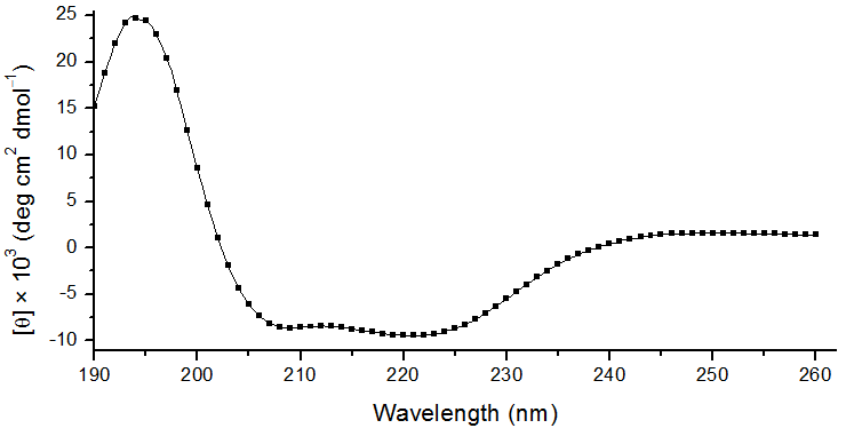


Figure 7.2: CD spectrophotometry of *Nme*Quin. Every second datum point is plotted.

structure of the proteins.

7.3 Kinetics

The enzyme activities and kinetic parameters for the mutant enzymes were measured (Table 7.3). Those for four of the proteins (F114A, F114R, R117K and R117A) were measured by Parker group member Dr Fiona Cochrane.

The mutation F139G had no negative effect on the reaction rate catalysed by *Nme*F139G, in fact the value of k_{cat} was slightly higher than for wild-type *Nme*KDO8PS. F139 is in the interface area between adjacent monomers and has no direct interactions with the active site in either its own subunit or that of the adjacent one (Figure 7.1a). However, K_m^{A5P} for this enzyme is raised relative to wild type, which suggests that the residue has some role to play in promoting the binding of A5P to the enzyme. The same effect on K_m^{A5P} was observed for *Nme*F114A, and there was a slight decrease in the value of k_{cat}. The mutation to Arg, which this residue is in DAH7PS, was more detrimental to the binding of PEP than the mutation to Ala, although K_m^{A5P} for *Nme*F114R is only one third that for *Nme*F114A.

Mutation of the interdigitating Arg to Lys (*Nme*R117K) reduced the value of k_{cat}, but also disrupted the binding of A5P, observed by a large increase in the value of K_m^{A5P}. The introduced Lys residue may partially fulfil the role of the native Arg residue, as in contrast, the mutation to Ala was devastating for enzyme activity, with this enzyme (*Nme*R117A) being barely active. Mutation to Gln, the form of this residue found in DAH7PS, resulted in a protein with a very large K_m^{A5P} value, but which was still active with the value of k_{cat} reduced to an eighth of that for wild-type *Nme*KDO8PS. These results suggest that the interdigitating Arg is important for the binding of A5P to the enzyme.

The double and triple mutant proteins, which all include mutation of R117, had dramatically increased K_m^{A5P} values. The k_{cat} values for the three enzymes were also much lower than for enzymes containing any of the single

Table 7.3: Kinetic parameters for interface-area mutants of *Nme*KDO8PS. Measurements for *Nme*F114A, *Nme*F114R, *Nme*R117A and *Nme*R117K were made by Dr Fiona Cochrane.

KDO8PS	K_m^{PEP} (µM)	K_m^{A5P} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{A5P} ($s^{-1}\,mM^{-1}$)
*Nme*WT	2.5 ± 0.2	12.0 ± 0.5	8.0 ± 0.1	660 ± 40
*Nme*F139G	14 ± 1	594 ± 30	8.6 ± 0.2	14 ± 1
*Nme*F114A	58 ± 6	873 ± 29	6.1 ± 0.2	7.0 ± 0.5
*Nme*F114R	95 ± 7	285 ± 18	3.0 ± 0.1	11 ± 1
*Nme*F114R/R117A		3700 ± 400	0.108 ± 0.007	0.029 ± 0.005
*Nme*F114R/R117Q		2742 ± 173	0.27 ± 0.01	0.10 ± 0.01
*Nme*F114R/R117Q/F139G		3600 ± 400	0.14 ± 0.01	0.039 ± 0.007
*Nme*R117K	22 ± 1	816 ± 32	4.8 ± 0.1	5.9 ± 0.4
*Nme*R117A[a]	–	–	–	–
*Nme*R117Q		3211 ± 196	1.06 ± 0.03	0.33 ± 0.03
*Nme*Quin[b]	–	–	–	–

[a] Barely active. Specific activity was 0.02 ± 0.01 U mg^{-1} in assays containing 200 µM PEP and 100 µM A5P.
[b] *Nme*Quin is *Nme*KPRS/F114R/R117Q/A126T/F139G/L7trun. No measurable activity.

mutations alone (except for *Nme*R117A, which was inactive). Interestingly, coupling mutation of F114R with R117A restored a small amount of activity compared to the devastating effect of the single R117A mutation. This shows that the complete loss of the Arg side chain could be tolerated by adding functionality by F114 substitution.

Assays for *Nme*Quin were performed at 25 °C because at 37 °C the protein precipitated. The precipitation caused a large increase in absorbance at the wavelength used for the continuous assay. It was also considered that precipitation may be detrimental to the activity of the enzyme. No activity was measurable for *Nme*Quin with A5P, E4P, R5P or 2dR5P.

Alternative substrates for A5P were also tested for *Nme*F114R/R117A, *Nme*F114R/R117Q and *Nme*F114R/R117Q/F139G. E4P, R5P and 2dR5P were not substrates for these three enzymes.

7.4 PEP binding to *Nme*Quin

As *Nme*Quin had no perceptible enzyme activity, an ITC experiment was performed to assess whether this was because it was no longer able to bind PEP. PEP bound to *Nme*Quin with a K_d value of 103 ± 7 µM (Figure 7.3). This value is 20 times greater than for wild-type *Nme*KDO8PS (K_D is 4.8 ± 0.2 µM). While PEP binding is compromised by the mutations, this result suggests that the reason for inactivity is more likely to be associated with the relative (in)ability to bind a phosphorylated aldose sugar substrate. This was not tested, but it may be possible to measure the enthalpy of binding of, for example, A5P or E4P also by ITC. These experiments could be performed in the presence of PEP, as the inactive enzyme will not turnover.

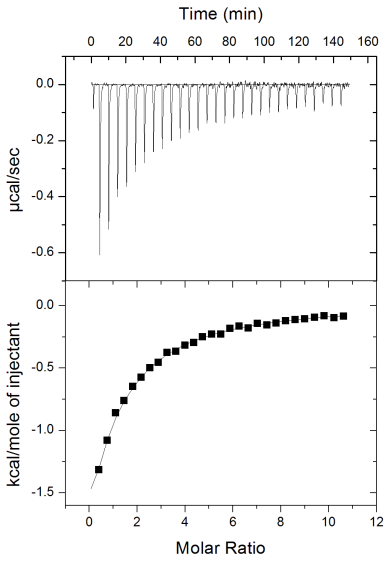


Figure 7.3: The interaction of PEP and *Nme*Quin, quantified by ITC.

7.5 Crystallography

Crystals were grown and diffraction data collected for eight mutants of *Nme*-KDO8PS: F114A, F114R, R117Q, R117K, F139G, F114R/R117A, F114R/R117Q and F114R/R117Q/F139G. Like for wild-type *Nme*KDO8PS, the asymmetric unit for each structure contains one complete tetramer. Generally the models for each structure were very similar to wild type and to each other, apart from the differences to the residues mutated.

In the structure of *Nme*F114A (collected by Dr Fiona Cochrane and Professor Geoff Jameson, Massey University), there are no major differences in the structure apart from the truncation of the side chain of position 114 caused by the mutation to Ala (RMSD of Cα atoms on the superimposed wild-type structure of 0.330 Å). The structure of *Nme*F114R (data set *a*) has little density for the side chain of F114R, suggesting that the Arg residue is in a disordered state. In chains B and C the side chain of F114R sits in a position sometimes occupied by an alternative conformation of F114 (Figure 7.4), and in chains A and D the F114R side chain is positioned pointing in toward the active site.

In a model from another dataset that was collected for *Nme*F114R (data set *b*), the side chains of F114R in all subunits are more disordered with density only resolved to the Cβ atom. Other than the side chains of F114R not being observed, there is one other notable change to the structure. For the first time ever in a structure of *Nme*KDO8PS, a fully ordered β7α7 loop is observed (in chain B only). The full β7α7 loop has only ever been observed before in structures of *Aae*KDO8PS in which both PEP and A5P (or the reaction intermediate or product) are bound. In this structure the β7α7 loop appears in a different conformation to that observed in *Aae*KDO8PS (Figure 7.5), representing the conformational flexibility in this dynamic and mobile loop. The conserved Gln (Gln202) at the bottom of the loop is in clearly defined density, and positioned pointing in toward the active site, in a similar orientation to which it is modelled in structures of *Aae*KDO8PS.

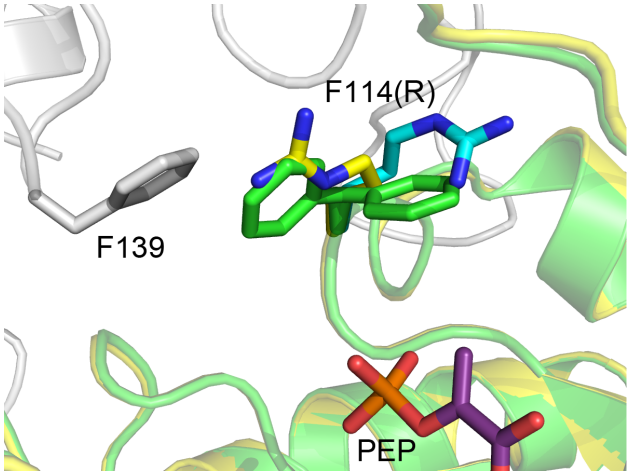


Figure 7.4: The structure of *Nme*F114R*a* (coloured yellow) superimposed onto wild-type *Nme*KDO8PS (PDB code 2QKF, coloured green). Chain D of wild-type *Nme*KDO8PS is shown where two observed rotamers of F114 are modelled. Chain D of *Nme*F114R*a* is shown, as is F114R from chain B (coloured cyan). The adjacent subunit is coloured white. PEP (carbon atoms coloured magenta) is shown from a superposition of the structure of *Aae*KDO8PS (PDB code 2NX3).

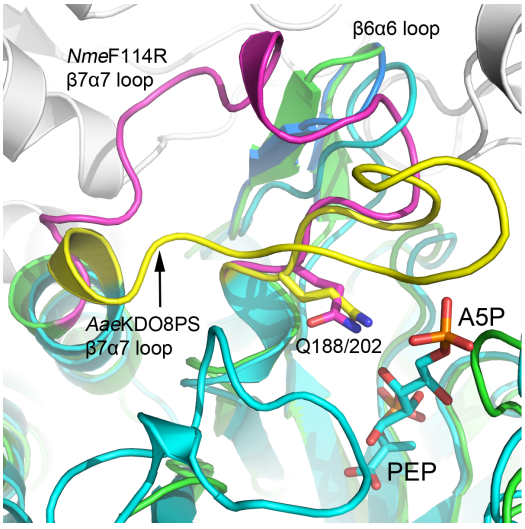


Figure 7.5: The structure of *Nme*F114R*b* (green) with that of *Aae*KDO8PS (PDB code 2NX3) superimposed (cyan). The β7α7 loop in *Nme*F114R*b* is coloured magenta, and in *Aae*KDO8PS is coloured yellow. The top of the β6α6 loop has shifted in the structure of *Nme*F114R; that of wild-type *Nme*KDO8PS is shown in cartoon and coloured blue. The other subunits of the *Nme*F114R*b* tetramer are coloured white.

Table 7.4: Crystal parameters, data collection, and refinement statistics for *Nme*F114A.

	*Nme*F114A
Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.35, 85.47, 163.06
Resolution range (Å)	39.95–1.90 (1.97–1.90)
No. of measurements	414454
No. of unique reflections	88194
Redundancy	4.7 (4.4)
Completeness (%)	97.8 (87.0)
$I/\sigma(I)$	8.8 (1.2)
R_{merge}	0.065 (0.601)
Wilson B value (Å^2)	38.3
Refinement	
Resolution (Å)	81.65–1.90
R_{cryst}	0.2507
R_{free}	0.2811
Amino acids (chain length of 280 residues)	248 + 252 + 247 + 256 residues; 7637 atom sites
No. of water molecules	307
No. of others	2 Cl^-, 1 Na^+
Mean B (Å^2)	
Protein	44.89
Water	37.85
Other	57.09
RMSD from target values	
Bond lengths (Å)	0.011
Bond angles (°)	1.427
Dihedral angles (°)	5.640
Ramachandran	
Most favoured (%)	92.3
Allowed (%)	6.6
Generously allowed (%)	1.0
Disallowed (%)	0.1

Table 7.5: Crystal parameters, data collection, and refinement statistics for *Nme*F114R*a* and *Nme*F114R*b*.

	*Nme*F114R*a*	*Nme*F114R*b*
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.80, 85.69, 163.64	81.86, 85.20, 163.36
Resolution range (Å)	47.98–1.80 (1.90–1.80)	39.70–1.80 (1.86–1.80)
No. of measurements	724402	440726
No. of unique reflections	107153	105281
Redundancy	6.8 (6.3)	4.2 (3.3)
Completeness (%)	100 (99.9)	98.7 (93.4)
$I/\sigma(I)$	6.1 (1.9)	12.6 (3.2)
R_{merge}	0.077 (0.397)	0.044 (0.302)
Wilson B value (Å^2)	22.7	31.8
	Refinement	
Resolution (Å)	46.03–1.80	81.65–1.80
R_{cryst}	0.2089	0.2241
R_{free}	0.2295	0.2457
Amino acids (chain length of 280 residues)	250 + 252 + 253 + 252 residues; 7813 atom sites	250 + 263 + 255 + 255 residues; 7889 atom sites
No. of water molecules	306	320
No. of others	–	1 glycerol
Mean B (Å^2)		
Protein	25.71	31.55
Water	25.31	30.92
Other	–	27.40
RMSD from target values		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	1.292	1.275
Dihedral angles (°)	5.166	5.424
Ramachandran		
Most favoured (%)	94.4	93.7
Allowed (%)	4.6	5.2
Generously allowed (%)	1.0	1.0
Disallowed (%)	0.0	0.1

The structure of *Nme*R117Q, where the interdigitating Arg residue is replaced by the Gln present in Iβ DAH7PS, like for the other structures, has no changes other than that for the mutation itself. In all chains the Gln occupies a similar position to the Arg in wild-type *Nme*KDO8PS (Figure 7.6a), although in chains B–D it is poorly ordered. When R117 is mutated to Lys, the mutation appears to have induced no other changes in structure, as observed in the model of *Nme*R117K (data collected by Dr Fiona Cochrane and Professor Geoff Jameson, Massey University). The Lys side chain occupies the same position as Arg in wild-type *Nme*KDO8PS (Figure 7.6a).

Mutation of F139 to Gly (the identity of this residue conserved in Iβ DAH7PS), has no effect on the observed conformation of F114, the residue with which the native Phe residue interacts across the subunit interface. However, the structure of *Nme*F139G reveals that the interdigitating Arg (R117) shifts toward the void created by the mutation in chains B and C (Figure 7.6b). In chain A, R117 is disordered and in chain D it occupies a similar position as in wild-type *Nme*KDO8PS. The increased mobility of

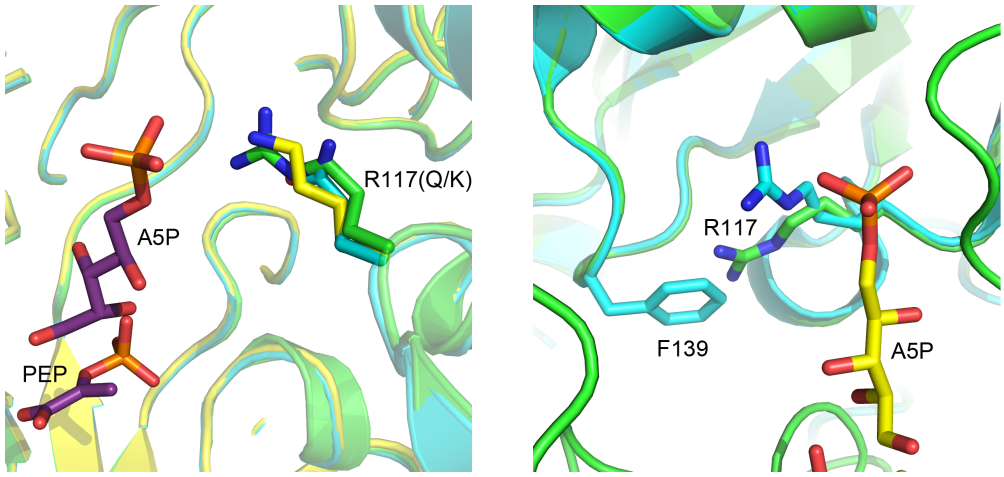


Figure 7.6: (a) The structure of *Nme*R117Q (cyan) superimposed with that of *Nme*R117K (yellow) and wild-type *Nme*KDO8PS (green, PDB code 2QKF). (b) The structure of *Nme*F139G (green) superimposed with that of wild-type *Nme*-KDO8PS (cyan). PEP and A5P [carbon atoms coloured magenta (a) or yellow (b)] are from the superimposed structure of *Aae*KDO8PS (PDB code 2NX3).

R117 in this model of *Nme*F139G correlates with the increase in $K_{\mathrm{m}}^{\mathrm{A5P}}$ for this enzyme.

In contrast to the structures of *Nme*F114R, where the side chain of F114R was generally disordered, in the structure of *Nme*F114R/R117A the side chain of F114R is well-ordered. In chains A and B, NE and NH2 of F114R hydrogen bond with the main-chain carbonyl of F114R and A113 respectively, from the adjacent subunit (Figure 7.7). Generally, in all chains, the side chain of F114R is close to encroaching upon the position occupied by R117 in structures of wild-type *Nme*KDO8PS. Hence, the truncation of R117 to Ala appears to have allowed room for a more favourable and ordered positioning of the side chain of F114R. However, the conformation F114R adopts in this structure is not that observed in structures of Iβ DAH7PS and thus it is not ideally positioned (unless it rearranges) to interact with the phosphate moiety of PEP, or the C2-OH of A5P or E4P.

In the structure of *Nme*F114R/R117Q the introduced Gln residue does appear to buttress F114R, pushing it further in toward the active site (Figure

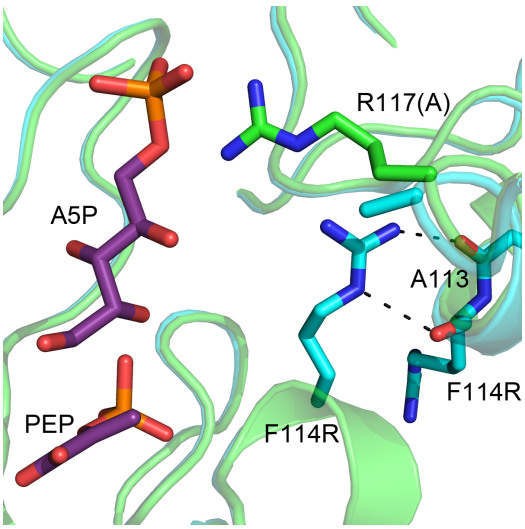


Figure 7.7: The structure of *Nme*F114R/R117A (cyan) superimposed with that of wild-type *Nme*KDO8PS (green, PDB code 2QKF). PEP and A5P (carbon atoms coloured magenta) are from the superimposed structure of *Aae*KDO8PS (PDB code 2NX3).

Table 7.6: Crystal parameters, data collection, and refinement statistics for *Nme*R117Q and *Nme*R117K.

	*Nme*R117Q	*Nme*R117K
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	81.47, 85.42, 162.71	81.53, 85.28, 162.92
Resolution range (Å)	45.79–1.86 (1.96–1.86)	39.93–1.75 (1.81–1.75)
No. of measurements	675001	462344
No. of unique reflections	95950	113642
Redundancy	7.0 (7.2)	4.1 (2.4)
Completeness (%)	99.9 (100)	98.8 (93.0)
$I/\sigma(I)$	5.1 (1.9)	14.1 (2.7)
R_{merge}	0.090 (0.398)	0.041 (0.325)
Wilson B value (Å^2)	23.5	32.2
	Refinement	
Resolution (Å)	41.31–1.86	81.51–1.75
R_{cryst}	0.2180	0.2320
R_{free}	0.2420	0.2516
Amino acids (chain length of 280 residues)	250 + 252 + 253 + 252 residues; 7811 atom sites	250 + 252 + 253 + 252 residues; 7807 atom sites
No. of water molecules	254	303
No. of others	–	–
Mean B (Å^2)		
Protein	28.02	32.97
Water	24.60	31.06
Other	–	–
RMSD from target values		
Bond lengths (Å)	0.009	0.008
Bond angles (°)	1.306	1.240
Dihedral angles (°)	5.407	5.116
Ramachandran		
Most favoured (%)	93.2	93.3
Allowed (%)	5.5	5.5
Generously allowed (%)	0.7	0.6
Disallowed (%)	0.6	0.6

Table 7.7: Crystal parameters, data collection, and refinement statistics for *Nme*F139G and *Nme*F114R/R117A.

	*Nme*F139G	*Nme*F114R/R117A
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	82.04, 85.94, 163.17	81.59, 85.37, 163.35
Resolution range (Å)	38.06–2.10 (2.18–2.10)	34.32–1.85 (1.95–1.85)
No. of measurements	195294	706400
No. of unique reflections	62633	95642
Redundancy	3.1 (3.1)	7.4 (7.1)
Completeness (%)	92.0 (99.9)	97.9 (96.8)
$I/\sigma(I)$	7.2 (2.3)	9.7 (1.9)
R_{merge}	0.080 (0.395)	0.066 (0.412)
Wilson B value (Å^2)	31.2	20.1
	Refinement	
Resolution (Å)	38.02–2.10	33.58–1.85
R_{cryst}	0.2357	0.1820
R_{free}	0.2657	0.2102
Amino acids (chain length of 280 residues)	250 + 252 + 253 + 252 residues; 7761 atom sites	250 + 251 + 253 + 252 residues; 7888 atom sites
No. of water molecules	86	910
No. of others	–	3 Cl^-, 1 Na^+
Mean B (Å^2)		
Protein	37.46	23.48
Water	27.95	33.04
Other	–	25.45
RMSD from target values		
Bond lengths (Å)	0.012	0.006
Bond angles (°)	1.504	1.111
Dihedral angles (°)	5.439	5.202
Ramachandran		
Most favoured (%)	93.8	93.3
Allowed (%)	5.1	5.4
Generously allowed (%)	0.5	0.7
Disallowed (%)	0.7	0.6

Table 7.8: Crystal parameters, data collection, and refinement statistics for *Nme*F114R/R117Q and *Nme*F114R/R117Q/F139G.

	*Nme*F114R/R117Q	*Nme*F114R/R117Q/ F139G
	Data collection	
Crystal system, space group	orthorhombic, $P2_12_12_1$	orthorhombic, $P2_12_12_1$
Unit cell parameters (Å)	82.15, 85.99, 163.11	81.67, 85.20, 162.85
Resolution range (Å)	48.00–2.00 (2.11–2.00)	41.20–1.75 (1.84–1.75)
No. of measurements	562713	822735
No. of unique reflections	78806	115042
Redundancy	7.1 (7.3)	7.2 (7.3)
Completeness (%)	100 (100)	100 (100)
$I/\sigma(I)$	4.1 (2.0)	4.5 (1.8)
R_{merge}	0.115 (0.377)	0.088 (0.419)
Wilson B value (Å^2)	21.5	21.7
	Refinement	
Resolution (Å)	45.93–2.00	40.82–1.75
R_{cryst}	0.1797	0.1944
R_{free}	0.2208	0.2221
Amino acids (chain length of 280 residues)	250 + 252 + 256 + 254 residues; 7929 atom sites	250 + 252 + 253 + 252 residues; 7777 atom sites
No. of water molecules	840	780
No. of others	5 Cl^-, 1 Na^+, 2 (1×0.5) glycerol	–
Mean B (Å^2)		
Protein	24.85	24.69
Water	35.12	35.07
Other	41.14	–
RMSD from target values		
Bond lengths (Å)	0.009	0.007
Bond angles (°)	1.088	1.020
Dihedral angles (°)	5.004	4.771
Ramachandran		
Most favoured (%)	92.7	93.1
Allowed (%)	6.0	5.8
Generously allowed (%)	0.7	0.5
Disallowed (%)	0.7	0.7

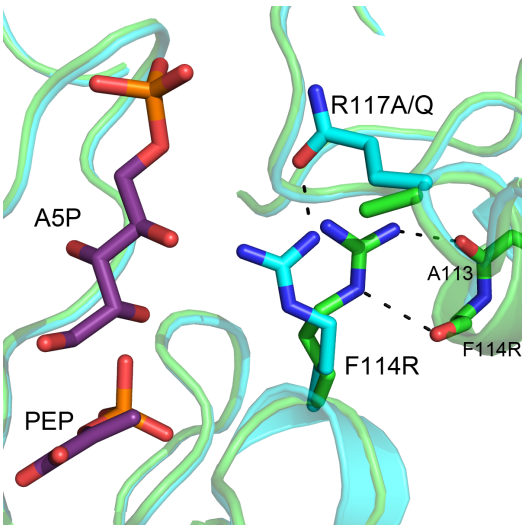


Figure 7.8: The structure of *Nme*F114R/R117Q (cyan) superimposed onto that of *Nme*F114R/R117A (green). PEP and A5P (carbon atoms coloured magenta) are from the superimposed structure of *Aae*KDO8PS (PDB code 2NX3).

7.8). In all four chains the side chain of F114R hydrogen bonds with R117Q and is in a closer position to the equivalent Arg in Iβ DAH7PS. In the structure of *Nme*F114R/R117Q/F139G the introduction of F139G has led to an increase in the range of conformations sampled by F114R. In chains A and D, F114R is in the same conformation as observed in the structure of *Nme*F114R/R117Q, and in chain B, F114R adopts a similar position to that observed in the structure of *Nme*F114R/R117A. In chain C the side chain of F114R is disordered.

Attempts to grow crystals of *Nme*Quin in the known *Nme*KDO8PS crystallisation conditions, which have produced diffracting crystals for all other mutants of *Nme*KDO8PS, were unsuccessful. The hanging drops incubated at 20 °C turned black within 24 h with a dark precipitate. Incubation at 5 °C produced a light amorphous precipitate in all drops. No further attempts were made to grow crystals of this protein.

7.6 Discussion

Residues of a conserved motif that contribute to the active site, and which are located in the interface area between adjacent monomers in KDO8PS, were mutated to assess their role in substrate binding. The mutant proteins (with the exception of *Nme*Quin) behaved similarly to wild-type *Nme*KDO8PS and were able to be purified using the same protocols. The melt temperatures measured for the proteins varied, an effect consistent with modification to the interface area between subunits. Structural analysis of the mutant proteins indicated that the introduced mutations had had little effect, and generally induced no major changes in structure. The mutations altered the kinetic parameters of all the mutant enzymes, with particular impact on the value of $K_{\mathrm{m}}^{\mathrm{A5P}}$ and k_{cat}.

7.6.1 Interdigitating Arg helps form A5P binding site

The equivalent residue to R117 is in *Aae*KDO8PS R106. In a previous study the *Aae*KDO8PS mutant R106G was created and the mutation was found to impair closure of the β7α7 loop.[79] It was suggested that the loss of the Arg residue affected the positioning of the β2α2 loop, which houses the KANR(S/T) motif that supplies many of the active-site interactions with A5P. It was suggested that the slight shift in the position of the β2α2 loop observed in crystal structures of *Aae*R106G (with no substrates bound, with only PEP bound, and with both PEP and A5P bound) negatively impacted the binding of A5P and was associated with the change in ability of the β7α7 loop to close. The $K_{\mathrm{m}}^{\mathrm{A5P}}$ value increased (from 7 ± 3 µM to 60 ± 20 µM for the *Aae*KDO8PS enzymes) and the changes in this value were attributed to both the direct loss of the Arg, which hydrogen bonds with the phosphate moiety of A5P, and indirectly, the changed dynamics of the β7α7 loop.

The results presented here are consistent with this picture and the role of the Arg in being to secure the formation of the complete binding site for the A5P phosphate moiety. Substitution of the Arg to Ala or Gln

(*Nme*R117A and *Nme*R117Q) both caused large increases in the value of $K_{\mathrm{m}}^{\mathrm{A5P}}$, whereas for *Nme*R117K, in which the Lys residue can fulfil some of the same role as the native Arg, the effect on $K_{\mathrm{m}}^{\mathrm{A5P}}$ was more modest. In contrast to the results for *Aae*KDO8PS, mutation of R117 in *Nme*KDO8PS appears to have no effect on the positioning of the β2α2 loop in structures of all *Nme*KDO8PS proteins in which it was mutated. All the *Nme*KDO8PS structures are however without any substrates bound. Additionally, the effect on $K_{\mathrm{m}}^{\mathrm{A5P}}$ were for *Nme*KDO8PS mutant enzymes much greater than that observed for *Aae*KDO8PS. This may be attributable to differences between mesophilic (*Nme*KDO8PS) and thermophilic (*Aae*KDO8PS) enzymes.

7.6.2 Subunit interface is important for substrate binding

F114 was mutated in *Nme*KDO8PS to Ala and Arg, the latter being the conserved identity of this residue in Iβ DAH7PS. In the crystal structures of these proteins there are no major changes observed other than for the mutations (except the unusual ordering of the β7α7 loop loop in one structure of *Nme*F114R). Despite this region of the protein normally being well-ordered, in the structure of *Nme*F114R the side chain of F114R is largely disordered. There is therefore limited structural insight into why these modifications to F114 have caused a modest increase in the values of $K_{\mathrm{m}}^{\mathrm{PEP}}$ and $K_{\mathrm{m}}^{\mathrm{A5P}}$ for these mutant enzymes.

Interestingly, the increase in value of $K_{\mathrm{m}}^{\mathrm{A5P}}$ for *Nme*F114R is more modest than for *Nme*F114A compared to wild-type *Nme*KDO8PS. This suggests that some steric bulk is useful in this space, but also that KDO8PS is relatively accommodating of an Arg residue at this position (although the melt temperature of *Nme*F114R is 3 ℃ to 5 ℃ less than wild-type *Nme*-KDO8PS in all conditions). This is perhaps not surprising given the similarity and evolutionary relationship between KDO8PS and DAH7PS. In structures of *Pfu*DAH7PS in which PEP is bound and E4P has been modelled, the equivalent Arg residue interacts not only with the phosphate moiety of PEP,

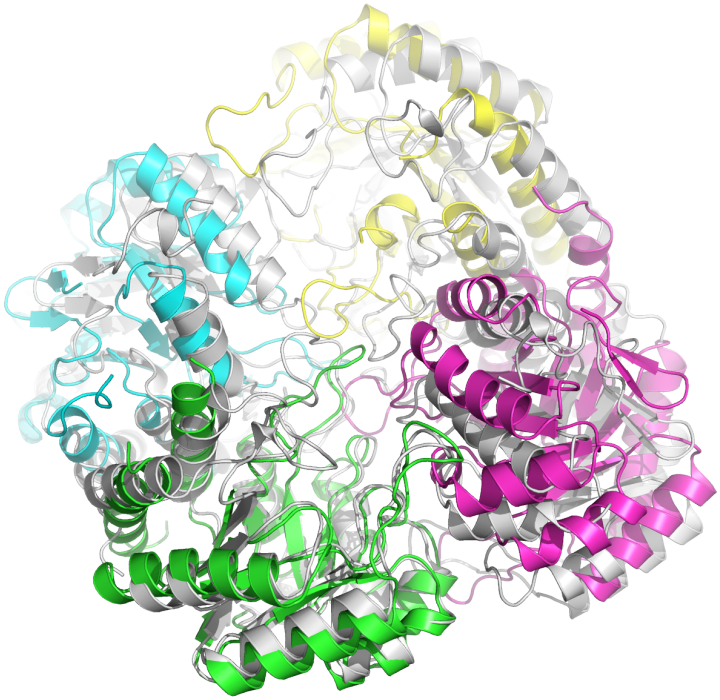
but with C2-OH of E4P (Figure 7.1b).[59] It is conceivable then that the Arg introduced into *Nme*KDO8PS by the mutation F114R may have a similar, but perhaps not optimised, interaction with C2-OH of A5P (and a better interaction with C2-OH of E4P, which has the opposite configuration at this stereogenic centre).

In contrast to the structures of *Nme*F114R, in the structures of the double and triple mutant proteins, in which the F114R mutation is coupled with mutation to R117 and F139 (*Nme*F114R/R117A, *Nme*F114R/R117Q and *Nme*F114R/R117Q/F139G), the side chain of F114R is ordered. Although removal of the interdigitating Arg functionality severely affects the ability of A5P to bind to the enzyme for reasons previously discussed, it does begin to allow F114R to be carefully positioned in the active site as would be desirable to optimise its interaction with the substrates. The conformation of the F114R side chain in structures of *Nme*KDO8PS is not in exactly the same position as in structures of Iβ DAH7PS (Figure 7.9b), however optimisation was observed in the buttressing of F114R by R117Q in comparison to the positions adopted by F114R when coupled with the mutation R117A.

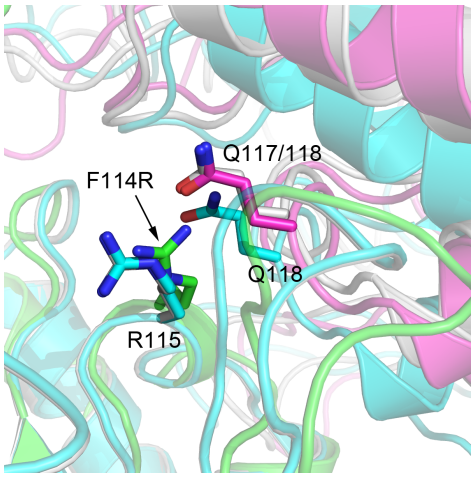
7.6.3 Effect of subunit packing

It is still difficult to understand why in KDO8PS position 114 (*Nme*KDO8PS numbering) is conserved as Phe, and in DAH7PS the equivalent residue is conserved as Arg. It seems reasonable that this difference is associated with other sequence differences, such as the composition of the motifs in which the residues are members, and also more generally the construction of the interface area between monomers. The results in this chapter demonstrate that the construction of this area is important for the binding of A5P. It would seem then that subunit packing and interface interactions are tightly coupled with active-site architecture, and hence substrate selection preferences. Indeed it was noted in Chapter 2 of this thesis that the interfaces between subunits are different between the two enzymes.

Careful comparison of the tetrameric structure of KDO8PS and Iβ



(a)



(b)

Figure 7.9: The structure of *Pfu*DAH7PS (PDB code 1ZCO) compared to the structure of *Nme*F114R/R117Q (chains independently coloured green, magenta, yellow and cyan). (a) Superposition constrained to chain A of both structures shows the relative difference in the arrangement of the other subunits. *Pfu*DAH7PS is coloured white. (b) Superposition of chain A of the *Pfu*DAH7PS tetramer (coloured cyan) on chain A on *Nme*F114R/R117Q, and the superposition of independent monomers of *Pfu*DAH7PS (coloured white) onto chain A and C of *Nme*F114R/R117Q.

DAH7PS reveals differences in the subunit assembly. The subunits of Iβ DAH7PS appear twisted relative to each other, relative to the subunits of KDO8PS (Figure 7.9a). Comparing the structure of *Nme*F114R/R117Q to that of *Pfu*DAH7PS, the position of F114R/R115 within each monomer is the same between both enzymes, however the position of the monomers relative to each other are different. Hence, in *Pfu*DAH7PS, the native Gln (Q118) is in a different position, compared to the equivalent R117Q in *Nme*KDO8PS, relative to R115/F114R in the adjacent monomer. However, superimposing independent monomers of *Pfu*DAH7PS onto *Nme*F114R/R117Q shows that the positions of the residues (including Gln) are all similar (Figure 7.9b). Although the conserved interface residues in KDO8PS can be converted to the equivalent residues of Iβ DAH7PS, unless the subunit assembly is also (somehow) changed, the mutated residues cannot be fulfilling exactly the same roles as in Iβ DAH7PS.

7.6.4 Role of PEP phosphate in reaction

It has been suggested that during the catalysed reaction the PEP phosphate could supply a hydrogen to the nascent negative charge on the carbonyl of A5P.[69] This supposes that the carbonyl is pointing toward the PEP phosphate rather than metal-binding site, and that the presence of Phe (as opposed to the Arg in DAH7PS) promotes a dianionic (rather than trianionic) form of PEP. However, the results with *Nme*F114R (and other enzymes containing this mutation) cast some doubt on this proposition. Though the mutant enzymes do have altered kinetic activity, with the mutation of F114 to Arg this area of the active site in *Nme*KDO8PS very much resembles that in Iβ DAH7PS.

7.6.5 Combining mutations does not convert enzyme function

That *Nme*Quin was inactive with A5P is not surprising: this protein combines several mutations that by themselves are extremely detrimental to (KDO8PS) enzyme activity. The inability of this enzyme to use E4P indicates that other changes must be necessary to successfully convert this KDO8PS into a DAH7PS. It is unfortunate then that the known *Nme*KDO8PS crystallisation conditions did not crystallise this protein, as access to structural information for this protein would aid the consideration and design of further changes. A notable difference between *Nme*Quin and DAH7PS is that *Nme*KDO8PS (and hence *Nme*Quin) is a metal-independent enzyme, yet all DAH7PSs are metal-dependent enzymes. Thus success may be more forthcoming in using either a metal-dependent enzyme as the scaffold on which to attempt the conversion of a KDO8PS into a DAH7PS, or alternatively by making *Nme*-KDO8PS a metal-dependent enzyme, as has previously been demonstrated to be possible.

7.7 Conclusion

In this chapter, nine mutant constructs that contained modifications to a subunit-interface area of *Nme*KDO8PS were characterised. The interface area was found to be an important contributor to the formation of the binding site for the substrate A5P. Changes to the conserved residues of the interface to make it more resemble that in DAH7PS were mostly tolerated by *Nme*KDO8PS, but caused negative changes in the kinetic parameters of the enzymes. Combining the interface-area mutations with other changes to the enzyme active site to create a DAH7PS-like protein resulted in a large relative decrease in protein stability, a loss of KDO8PS activity, and no gain of DAH7PS activity.

Chapter 8

Summary of thesis and overall conclusions

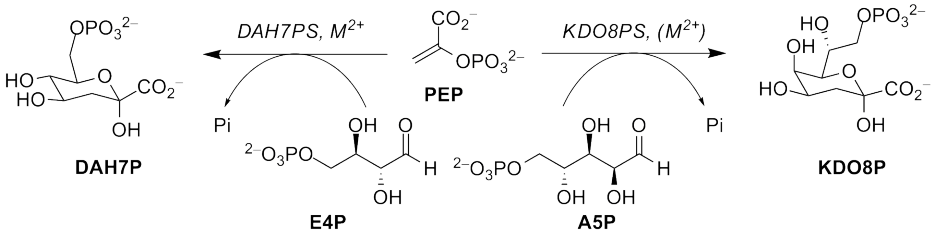
In this thesis, the relationship between sequence, structure and function in the enzyme KDO8PS was explored. Sequence analysis of KDO8PS and the related Iβ DAH7PS enzymes highlighted conserved residues, and in combination with structural analysis, the important sequence features, which are directly responsible for the active-site architecture and function, were identified. In addition, the sequence features important for the selection of the phosphorylated aldose substrate were pinpointed. These conserved active-site residues were systematically investigated by mutational analysis in a metal-dependent and metal-independent KDO8PS. This investigation included mutating the residues and motifs to the equivalent residues of the closely-related enzyme DAH7PS, in an attempt to understand sequence differences between KDO8PS and Iβ DAH7PS. The individual sequence-feature differences were combined to engineer a 'KDO8PS' enzyme with an active site that resembles a typical Iβ DAH7PS. However, this engineered protein was catalytically inactive for use of all phosphorylated aldoses tested, despite still being able to bind PEP.

Investigation of the individual sequence features of KDO8PS related

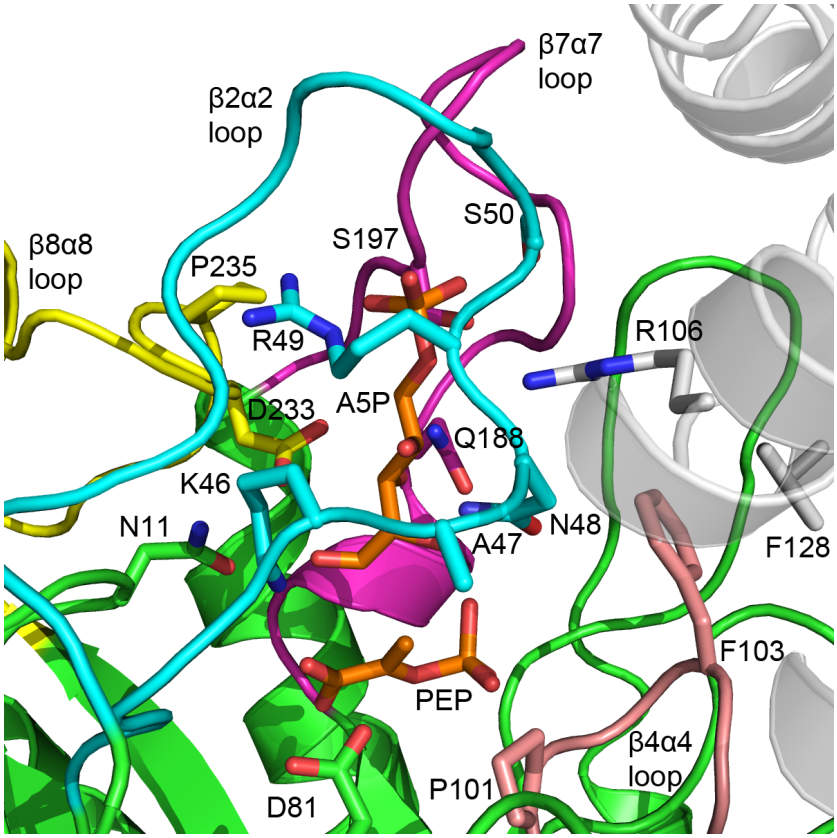
to aldose substrate selection did uncover information about their roles in the enzyme catalysed reaction. In KDO8PS there is a conserved motif on the β2α2 loop, KANR(S/T), whereas in Iβ DAH7PS the equivalent motif in sequence and space is KPR(S/T). The Lys, Arg, Ser/Thr all play similar roles in the two enzymes. The Lys was critical for catalysis in KDO8PS, suggesting a role intimately involved in a proton relay chain, while the Arg and Ser/Thr likely interact with the phosphate moiety of A5P or E4P. In KDO8PS, where the Pro of the Iβ DAH7PS motif is replaced with the residues AlaAsn, the Asn residue has hydrogen-bonding contacts with the hydroxyl group on C2 of A5P. This interaction was found to be important for helping to arrange A5P in the active site and removal of this residue (by substitution to Ala) lowered substrate specificity.

The extended (and mobile) β7α7 loop in KDO8PS was also investigated. This loop features two conserved residues that interact with A5P. A conserved Ser residue near the top of the loop interacts with the phosphate moiety of A5P, and a conserved Gln interacts with A5P through its C3-OH. The β7α7 loop was truncated to resemble the much shorter length of this loop in Iβ DAH7PS, and the conserved Ser and Gln were individually, and in combination, mutated to Ala. The kinetic behaviour of the mutant enzymes indicated that the presence of the loop as a whole is more important for efficient catalysis than the interactions provided by the Ser and Gln residues.

The role of a conserved Asp residue in the binding pocket of PEP in KDO8PS was explored. In Iβ DAH7PS, the equivalent residue is conserved as a Glu, which offers the same functionality but has a side chain one-carbon longer. Although the residue in neither enzyme has direct interactions with either substrate, mutation of the Asp in KDO8PS affected the binding of both substrates and the catalytic rate. While no definitive reason for the difference in conserved residue between the two enzymes could be elucidated, the difference, from structural analysis, seems to correlate with the identity of a residue on the nearby β4α4 loop. In Iβ DAH7PS, a Gly is present on the β4α4 loop which allows the native Glu to be accommodated in a conformation inaccessible in KDO8PS because of the presence of a Pro in the same position



(a) The reactions catalysed by DAH7PS and KDO8PS



(b) The structure of *Aae*KDO8PS (PDB code 2NX3). The equivalent residues in *Aae*KDO8PS to those investigated in this thesis in *Afe*KDO8PS and *Nme*KDO8PS are shown as sticks. The β2α2 loop is coloured cyan, the β4α4 loop is coloured salmon, the β7α7 loop is coloured magenta, the β8α8 loop is coloured yellow, and the carbon atoms of PEP and A5P are coloured orange. The adjacent subunit, contributing the Arg to the A5P-phosphate binding site, is coloured white.

Figure 8.1

as the Gly. Coupling the Asp-to-Glu mutation with that of Pro-to-Gly may reduce the disruption to substrate binding caused by the introduction of the Glu residue.

The final sequence feature investigated was the PAFLxR motif in KDO-8PS. The Arg of the motif, which interdigitates into the active site from the adjacent subunit, had previously been identified as having an important role to play in securing the formation of the binding site for the phosphate moiety of A5P.[79] The results of the studies in this thesis also support this role. In Iβ DAH7PS the equivalent motif is GARNxQ. Instead of interdigitating, the Gln of the Iβ DAH7PS motif buttresses the position of the Arg in the adjacent monomer. The Arg of the motif in Iβ DAH7PS interacts with the phosphate moiety of PEP and likely also interacts with the C2-OH of E4P. When in *Nme*KDO8PS the PAFLxR motif was mutated to PARLxQ, to resemble Iβ DAH7PS, the Gln was similarly observed to buttress the position of the Arg. It was also observed that KDO8PS and Iβ DAH7PS display, at least in crystal structures, slightly different monomer packing. Although in individual monomers the investigated residues were similarly placed, the positions relative to each other in the tetrameric structure were slightly different. This means that the residues in this interface area, mutated to resemble those in Iβ DAH7PS, will not play exactly the same roles as they do in Iβ DAH7PS, as the positions and therefore interactions they have depend on the arrangement of the subunits in the tetrameric structure.

For KDO8PS, because of the interdigitating Arg residue, a dimer appears to be the minimum quaternary structure required for function. In Iβ DAH-7PSs there are no residues that interdigitate from adjacent subunits, and in type II DAH7PSs (such as *Mtu*DAH7PS[54]), the subunit interface belonging to the putatively required dimer in KDO8PS (and the specific buttressing it provides) does not exist due to the arrangement of the subunits in the tetramer. Therefore, in contrast to KDO8PS, it would appear that for DAH-7PS the minimal functional unit is a monomer. Experiments to disrupt the quaternary structure in both KDO8PS and DAH7PS could be used to investigate its functional importance.

8.1 Differences between enzymes is more complex than first envisioned

Iβ DAH7PS and KDO8PS appear on many levels to be very similar enzymes. The active site of each enzyme is very similar to that of the other, and many residues are equally conserved between both enzymes. The metal-binding site, PEP-binding site, and in many respects, the binding sites for A5P and E4P are very similar. Given the similarity and evolutionary relationship of the two enzymes, it should be possible, as was partly the aim in this thesis, to reverse the process of evolution and convert a KDO8PS into an Iβ DAH7PS. To do this requires a change in the preference for only the phosphorylated aldose: to again accept a four-carbon sugar, with less preference for the configuration of the C2-OH.

The results presented in this thesis suggest that modifying the most obvious differences that would give rise to the desired different substrate specificity, is not an efficacious strategy, at least for the *Nme*KDO8PS template. The obvious differences, such as KANR(S/T)/KPR(S/T), β7α7 loop truncation, GARNxQ/PAFLxR are likely still important to enable successful conversion, but these differences may well be dependent on other less evident differences between the two enzymes. The sequence and structural analysis in Chapter 2 highlighted that there are many other conserved differences between the two enzymes. For example, it was observed that the β2α2 loop [which houses the KANR(S/T)/KPR(S/T) motifs] is buttressed quite differently across the monomer interface in the two enzymes.

An important distinction between KDO8PS and DAH7PS is metal dependency. The metal ion in DAH7PS appears to have an important and indispensable role to play, likely to be responsible for directly coordinating to the aldehyde group of E4P.[69] In DAH7PS, it has not been possible to create metal-independent forms by the same mutation of the metal-ligand Cys to Asn, which, generally, enables interconversion in KDO8PSs.

The template used in this thesis on which the identified changes were

combined, was *Nme*KDO8PS, which is a natively metal-independent enzyme. *Nme*KDO8PS was chosen, in particular preference over using *Afe*KDO8PS (the other KDO8PS studied in this thesis) because crystallisation conditions were known and there was therefore access to useful structural information. The metal-dependent *Afe*KDO8PS also has some sequence features more common to metal-independent KDO8PSs, which disfavoured its use.

A more successful strategy to achieve conversion of substrate specificity may be to use a metal-dependent KDO8PS as the template on which to assemble the multiple changes. The two mutations (N23C/C246S) that give rise to obligate metal dependency for *Nme*KDO8PS could be introduced, or another metal-dependent KDO8PS could be used. In hindsight, given that crystals of *Nme*Quin did not grow in the known *Nme*KDO8PS crystallisation conditions, there is little advantage in using *Nme*KDO8PS as the template, even if engineered metal-dependency was also included. *Aae*KDO8PS is a well-studied metal-dependent KDO8PS, with accessible crystallisation conditions.[24,40,45,67,78,79,82] The thermophilicity of *Aae*KDO8PS, and compatibility of the crystallisation conditions to ligand binding, has meant that it has been possible to determine substrate-bound structures of the enzyme.[24,27,72] Phylogenetic analysis has suggested that *Aae*KDO8PS is more closely related than many other KDO8PS to the theorised ancestral Iβ DAH7PS-like protein.[64] These factors suggest it may be ideal to use *Aae*KDO8PS as the template on which to engineer the changes identified in this thesis.

For all KDO8PSs, the sequence features most obviously associated with selection of the sugar substrate are conserved: an extended β7α7 loop, KANR(S/T) motif on the β2α2 loop, and PAFLxR motif. The equivalent sequence features in Iβ DAH7PS are also conserved, but as different residues. However, between the two enzymes there are also other conserved sequence differences less obviously associated with substrate specificity. Future studies will need to distinguish the less obvious differences that are truly associated with substrate specificity from those that are not. Perhaps too, this indicates that the two enzymes are not so similar despite the shared active-site architectures; the difficulty in converting *Nme*KDO8PS to become DAH7PS-like

hints that the key differences between the enzymes are more complex than first envisioned. To investigate these differences between the two enzymes, an alternate approach to rational redesign is an option (*vide infra*).

8.2 Approaches toward evolving enzyme function

8.2.1 Natural evolution

The TIM-barrel is the most common fold of structurally characterised proteins,[98] but it is unclear whether this dominance is the result of divergent evolution, or from convergent evolution (or both) to what is in many ways an exemplary fold.[99] What is clear however is that this fold is an adaptable scaffold on which a diverse number of functionally different enzymes have evolved. The active sites of all TIM-barrel enzymes are positioned, independent of structure, at the C-terminal ends of the β-strands.[100] The loops that extend the ends of the β-strands and connect with the subsequent α-helices, provide the functional groups to the active sites. In the evolution of new functions, some of these groups are retained while some are altered.[100]

The TIM-barrel proteins are split into a number of different superfamilies. One theory is that the core $(\beta/\alpha)_8$ barrel originated from combinations of $(\beta/\alpha)_{2N}$ subdomains, and different assemblies of these have formed the ancestors of the different superfamilies.[99] From these original combinations, gene duplication and mutation has given rise to the large variety of constituent enzymes in each family.

Two scenarios exist for when gene duplication occurs.[99] Most likely, the new copy of the gene 'accidentally' catalyses a new reaction, through a promiscuity of activity that is apparently widespread. Alternatively, the duplicated gene is not capable of a new function immediately, but with a limited number of alterations will become capable. There are however no

known examples of this second scenario.

More specifically, the adaptation of new function can come via three routes.[101] Firstly, preservation of chemical mechanism, where the new function has different substrate specificity. Secondly, when ligand specificity is dominant, the old scaffold is useful as it provides binding for a desired ligand. Or third, a more opportunistic route, where the active-site architecture is dominant and functional groups useful for a new function are exploited and adapted. The families 'enolase C-terminal domain-like' and 'phosphoenolpyruvate/pyruvate domain' are, respectively, examples of the first and second route.[99] The orotidine 5'-monophosphate decarboxylate superfamily members have different mechanisms yet homologous sequences and structures, and are an example of the third route.[99]

Aldolase superfamily

KDO8PS and DAH7PS belong to the aldolase superfamily. This diverse superfamily consists of mechanistically distinct enzymes, which all catalyse aldol reactions. The superfamily includes class I aldolases (Schiff base stabilises an enediolate intermediate), some class II aldolases (divalent metal ion stabilises an enediolate intermediate), porphobilinogen synthase (Schiff base intermediate), and lastly those enzymes which use PEP as an enolpyruvate source for an aldol reaction.[99] This last subfamily contains the homologous heptulosonate (DAH7PS), octulosonate (KDO8PS) and neuraminate phosphate synthases. However, not all aldolases with a TIM-barrel fold are in this superfamily. Some are members of the 'ribulose-phosphate binding barrel' and 'phosphoenolpyruvate/pyruvate domain' superfamilies. The latter includes other class II aldolases, and those that use enolpyruvate as an intermediate.

There is also considerable structural diversity within the aldolase superfamily.[99] Members show no greater structural homology to each other than to enzymes in the 'ribulose-phosphate binding barrel' and 'phosphoenolpyruvate/pyruvate domain' superfamilies. Therefore some uncertainty remains as to whether these diverse enzymes actually do constitute an 'aldolase'

superfamily.[99]

In the specific case of KDO8PS and DAH7PS however, it is reasonable to suggest these enzymes represent a case of divergent evolution from a substrate-promiscuous ancestor, where both mechanism and, largely, substrate specificity is the same, with evolution of only the specificity for the phosphorylated aldose sugar.

8.2.2 Rational redesign and directed evolution

The evolution of function in TIM-barrel enzymes, specifically the alteration of substrate specificity, has been successfully performed in the laboratory environment, both by rational redesign (as was attempted in this thesis), directed evolution, and combinations of the two techniques.

Directed evolution provides a means to engineer new functions in proteins using artificial selection. Experiments typically include a diversification step, where random mutations or recombinations are made to create a large library of variants (for example by error-prone PCR, or deoxyribonucleic acid (DNA) shuffling), before a selection step where variants with the desired property are either screened or selected. Finally, those with the desired property are analysed (for example by DNA sequencing). These three steps are a single 'round'. Directed evolution experiments will often involve several rounds, using the selections from the previous round as the input for the next round.

- Substrate specificity of members of the enolase superfamily of enzymes has been changed, both by rational engineering and by directed evolution. The enolase superfamily enzymes have a $(\beta/\alpha)_8$-barrel structure, which houses the catalytic residues, and an N-terminal α and β capping domain that determines the substrate specificity. With minimal base changes that produce one, two, or three amino acid changes, a change in substrate selection was achieved and tuned in an additive fashion, to turn an L-Ala-D/L-Glu epimerase into an *o*-succinylbenzoate synthase, an enolase from a different subfamily.[102–104] Likewise, the same new activity was

also able to be added to muconate lactonising enzyme II, from a different subfamily to the epimerase, by directed evolution experiments.[102]

- Rational redesign has also been successfully used to convert 3α-hydroxysteroid dehydrogenase into a 20α-hydroxysteroid dehydrogenase using loop chimeras.[105] Interchanging the whole loops was necessary, as individual mutations to the critical loops was not successful in modifying the specificity.

- The enzyme D-hydantoinase is a $(\beta/\alpha)_8$-barrel enzyme that has gating loops forming the substrate-binding pocket. Rational manipulation of these loops, by single, double, and triple mutations, was able to induce distinct changes in substrate specificity.[106]

- The enzymes HisA and TrpF catalyse similar reactions in the biosynthesis of histidine and tryptophan respectively, catalysing an Amadori rearrangement in the isomerisation of an aminoaldose to an aminoketose. Both enzymes possess a similar $(\beta/\alpha)_8$-barrel structure, and it is believed that both evolved from a common ancestral enzyme, which had low substrate specificity. Using directed evolution, the process of natural evolution was able to be reversed for a HisA, thus recreating an enzyme resembling the ancestral protein with promiscuous substrate selection.[107] The combination of mutations that gave rise to the change in specificity were not all in close proximity to the active site.

- The $(\beta/\alpha)_8$-barrel enzyme D-sialic acid aldolase, was converted to become an efficient L-3-deoxy-*manno*-2-octulosonic acid aldolase by directed evolution.[108] The evolved enzyme had eight amino acid changes, all occurring outside the active site.

- A change in substrate specificity was also achieved using directed evolution for the *E. coli* 2-keto-3-deoxy-6-phosphogluconate aldolase.[109] A double mutant variant of the enzyme had an altered substrate profile, caused by an essential lysine shifting to an adjacent β-strand.

In many cases it seems that a small number of amino acid substitutions

for enzymes with a $(\beta/\alpha)_8$-barrel structure can significantly affect substrate specificity. Furthermore, the altered amino acids need not necessarily be close to the active site. This contrasts with KDO8PS and Iβ DAH7PS, where a number of fundamental differences were identified between the enzymes, and that none of these alone was able to sufficiently alter substrate specificity. That amino acids at sites remote to the active site can have a role in substrate selection for other enzymes of a similar fold does however correlate with many of the identified conserved differences between KDO8PS and Iβ DAH7PS being not obviously associated with substrate selection.

Given the success with which the directed evolution approach has had in changing substrate specificity for other $(\beta/\alpha)_8$-barrel enzymes, it may also work in this context of evolving a KDO8PS to become a DAH7PS. Indeed, previously, attempts have been made to use directed evolution to modify *Porphyromonas gingivalis* KDO8PS to gain DAH7PS activity, however these were unsuccessful.[110] DAH7PS activity has been successfully evolved, using directed evolution, on a 2-keto-3-deoxy-6-phosphogalactonate aldolase scaffold, using a similar approach.[111] These attempts both used a strain of *E. coli* with knockouts of the three chromosomal DAH7PS isozyme genes. These strains grow on minimal media supplemented with the aromatic amino acids and some other vitamins. Modulating the concentration of these supplements allows tuning of the selection system. Transformation with plasmids containing the gene for a putative DAH7PS, with expression under the control of a inducible *lac* operon, allows selection for mutated constructs with the desired DAH7PS activity.

Perhaps, the ideal strategy may be a combination of rational engineering and directed evolution: to start with a *Nme*Quin-like template, on which the obvious changes are already engineered to a metal-dependent KDO8PS, and then use directed evolution to uncover the less obvious changes necessary to gain DAH7PS function. Alternatively, an opposite strategy could be employed to uncover the key determinants of substrate specificity in both enzymes. Using a broadly promiscuous Iβ DAH7PS as a starting point, the minimum changes necessary to create a KDO8PS from a DAH7PS could be investigated.

This thesis has demonstrated that the distinguishing features of KDO-8PS, as compared to Iβ DAH7PS, are intricate and complex. Contemporary KDO8PS enzymes have many more conserved differences to DAH7PS than would allow a simple switch of substrate specificity to exist. It is apparent that the $(\beta/\alpha)_8$ barrel is an adaptable protein fold on which new enzyme functions can evolve, and this holds promise for the potential of using it as a scaffold to engineer novel enzyme functions and properties.

Chapter 9

Experimental procedures

9.1 General methods

Protein structure figures

Figures of protein structures were created using PyMOL (version 1.5, Schrödinger LLC).[112]

Multiple sequence alignments

Multiple sequence alignments were generated using ClustalX (version 2.0).[113]

Water

All water used in these experiments was from a Millipore Milli-Q system. This water is referred to as Milli-Q water. Water used for PCRs was autoclaved prior to use.

pH measurements

The pH of solutions was measured using a Denver Instruments UB-10 Ultra-Basic pH meter, with either a standard or micro-probe. Solutions were made more acidic by addition of 1 M or 10 M HCl (or when specifically required, glacial acetic acid) and basic by addition of either 1 M or 10 M NaOH.

Removal of metal ions from solutions

Divalent metal ions were removed from solutions by treatment with Chelex® 100 resin (Bio-Rad). The resin was added to a solution and then stirred for at least 1 h before the resin was removed by filtration (0.2 µm). The resin invariably caused the pH of the solution to increase, sometimes necessitating adjustment to correct the pH of any resin-treated solution.

Glycerol stocks

All plasmid-containing strains of *E. coli* were stored frozen at −80 °C as glycerol stocks. Glycerol stocks were created by mixing 0.4 mL of 50 % (v/v) glycerol with 0.8 mL of an overnight culture in either a 1.6 mL micro-centrifuge or screw-top tube.

Antibiotics

The genes for *Nme*KDO8PS and *Afe*KDO8PS had previously been cloned into the pT7-7 vector,[47,81] which carries resistance to ampicillin. Ampicillin was added to all growth media to a final concentration of 0.1 mg mL^{-1}. Stock solutions of ampicillin at 100 mg mL^{-1} were created, filter sterilised and stored at −80 °C, and freeze-thawed a maximum of two times.

Site-directed mutagenesis

Mutagenesis was performed using either a Quikchange® Lightning Mutagenesis Kit (Stratagene), Quikchange® XL Mutagenesis Kit, or with Pwo SuperYield DNA Polymerase (Roche), or Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific). PCRs were performed in an Veriti® 96-well Thermal Cycler (Applied Biosystems) or an iCycler (Bio-Rad). Typically each PCR was 25 µL and used 25 ng of plasmid template and 62 ng of each primer. For 50 µL reactions the amounts of plasmid and primers were doubled. Cycling protocols were those recommended by each kit or polymerase. Primers for mutagenesis were synthesised by GeneWorks or Invitrogen and resuspended in either TE buffer (10 mM Tris-HCl, 0.1 mM EDTA) or sterilised Milli-Q water to a concentration of 300 µM. Template DNA (which is methylated) was digested by treatment with the restriction enzyme DpnI according to the protocols of the kit, Invitrogen, or Roche, depending on the source of the enzyme.

Agarose gel electrophoresis

Separation of DNA was performed by agarose gel electrophoresis, using either 1 % (w/v) agarose gels or pre-cast E-Gel® 1.2 % (w/v) agarose gels (Invitrogen). DNA was stained for both self-poured and pre-cast gels by inclusion of SYBR® Safe DNA stain.

Agarose gels were created by dissolving 1 % (w/v) agarose in 30 mL of Tris-acetate-EDTA (TAE) buffer and heating in a microwave oven until the agarose was dissolved. The hot agarose solution was left to cool in a 60 ℃ oven. Immediately before pouring, SYBR® Safe DNA stain was added. Samples were mixed with a (6×) sample loading dye. Electrophoresis was performed for 50 min at 85 V in TAE buffer using a Mini-Sub® Cell GT (Bio-Rad).

TAE buffer: 50 mM Tris, 1 mM EDTA, 20 mM acetic acid

Sample loading dye (6×): 60 mM Tris-HCl, 60 mM EDTA, 0.2 % (w/v) o-

range G, 0.05 % (w/v) xylene cyanol ff, 60 % (v/v) glycerol

The E-Gel® agarose gels were used according to the manufacturer's instructions. Samples were loaded without a sample loading dye directly into the wells, and separation could be monitored in real-time during electrophoresis using the E-Gel® iBase™ Safe Imager™ systems at 480 nm.

Typically 15 µL from 25 µL PCR reactions and 20 µL from 50 µL PCR reactions were loaded onto the agarose gels for visualisation of amplification products. To one well of each agarose gel, 2 µL of DNA ladder (1 kb Plus DNA Ladder [Invitrogen] or 1 kb DNA Ladder [New England BioLabs]) was used to estimate DNA size. Stained DNA was visualised and photographed under UV (302 nm) using a Molecular Imager® Gel Doc™ XR (Bio-Rad).

Plasmid extraction and purification

Plasmids were extracted using a High Pure Plasmid Isolation Kit (Roche). Small cultures (5 mL) of plasmid-containing cells were grown overnight with the appropriate antibiotic(s). The cultures were harvested the following morning and plasmids extracted and purified as per the kit instructions, with two variations. Firstly, as optionally specified in the kit instructions, the glass fibres were pre-equilibrated with binding buffer to increase plasmid yield, and secondly, plasmids were eluted from the fibres using 70 µL to 100 µL rather than 100 µL of elution buffer, to ensure a sufficient plasmid concentration for downstream uses.

The concentration of purified plasmid was measured by absorption at 260 nm using an appropriately blanked Nanodrop ND-1 000 spectrophotometer.

DNA sequencing

All DNA sequencing was performed at Canterbury Sequencing and Genotyping using an Applied Biosystems 3130xl Genetic Analyzer with BigDye Terminator

v3.1 (Applied Biosystems) sequencing chemistry. Typically 200 ng to 250 ng of purified double-stranded plasmid and 3.2 µM of primer were supplied for each sample. As all KDO8PS genes were in the pT7-7 vector the same forward and reverse primers were used for all sequencing.

Forward: 5' `TAA TAC GAC TCA CTA TAG GGA GA`
Reverse: 5' `GTT TAC TCA TAT ATA CTT TAG AT`

Preparation of chemically competent cells

Chemically competent cells were prepared by two similar methods.

In the first method,[114] an overnight culture of the cells to be made competent was grown in the presence of the appropriate antibiotic(s). The following day lysogeny broth (LB) medium with the appropriate antibiotic was inoculated with one hundredth the volume of the overnight culture. Cells were grown at 37 °C until reaching OD_{600} of 0.3 AU to 0.6 AU. The culture medium was then incubated on ice for 10 min to 15 min, before being harvested by centrifugation for 15 min at 4 000 rpm and 4 °C, in a pre-chilled centrifuge and rotor. The pellet was then resuspended in 18 mL of RF1 solution and incubated on ice for 30 min. The cells were then harvested again by the same centrifugation method, and resuspended in 4 mL of RF2 solution. Suspended cells were then split into 50 µL to 100 µL aliquots, flash-frozen in liquid nitrogen and stored at −80 °C.

RF1 solution: 100 mM $RbCl_2$, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$, pH 5.8 (adjusted using concentrated acetic acid).

RF2 solution: 10 mM $RbCl_2$, 10 mM MOPS, 75 mM $CaCl_2 \cdot 6\,H_2O$, 15 % (v/v) glycerol, pH 5.8 (adjusted using NaOH).

The RF1 and RF2 solutions were filter-sterilised (0.2 µm) rather than autoclaved to avoid formation of a Rb precipitate, and stored at 4 °C.

In the second method to prepare competent cells, an overnight culture was grown including the appropriate antibiotic(s). The next morning, a

100 mL culture was inoculated with 2 mL of the overnight culture, and grown at 37 ℃ until OD_{600} reached 0.3 AU to 0.4 AU. At this point the cells were transferred to a sterile 50 mL centrifuge tube, and centrifuged at 4 000 rpm for 15 min at 4 ℃ to pellet the cells. The supernatant liquid was poured off and the pellet gently resuspended on ice in 50 mL of freshly-made sterile 100 mM $CaCl_2$. The resuspended cells were left on ice for 30 min, after which they were again centrifuged, and the pellet gently resuspended in 10 mL of cold $CaCl_2$ on ice. The cells were then finally incubated on ice for at least 2 h, before being gently mixed with pre-chilled sterile 60 % (v/v) glycerol to a final concentration of 15 % (v/v). The cells were then split into aliquots, snap-frozen in liquid nitrogen, and stored at −80 ℃.

Transformation

Chemically competent cells (50 µL to 100 µL aliquots) were thawed on ice before addition of 2 µL of either PCR product or purified plasmid. The mixture was left on ice for 30 min then placed in a 42 ℃ water bath for 30 s to 45 s before being placed on ice again for 2 min to 5 min. Cells were then outgrown for 1 h at 37 ℃ while shaking after addition of 150 µL to 200 µL super optimal broth (SOC) medium.

A portion (50 µL to 200 µL) of the transformed cells was spread directly onto an LB-agar plate containing the appropriate antibiotic(s) and left to grow on the inverted medium overnight at 37 ℃.

SOC medium: 2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 20 mM glucose. Filter sterilised in aliquots.

Cell lines

The cell lines used for plasmid propagation were either *E. coli* XL1-Blue or TOP10 (Invitrogen). For protein expression either *E. coli* BL21 (DE3) or

BL21 (DE3) Star (Invitrogen) cell lines were used.

Culture media

LB-agar was prepared by dissolving either LB-agar (Miller's) ($37\,\mathrm{g\,L^{-1}}$) or LB (Lennox L) ($20\,\mathrm{g\,L^{-1}}$) and agar ($15\,\mathrm{g\,L^{-1}}$) base in Milli-Q water and sterilised by autoclaving. The LB-agar solution was then heated in a microwave oven until boiling, and left to cool in a 60 °C oven. The LB-agar solution was then further cooled to $\lesssim$50 °C before antibiotic(s) were added immediately prior to pouring into round petri dishes.

LB media for plasmid extraction, pre-cultures and protein expression cultures were prepared by dissolving $20\,\mathrm{g\,L^{-1}}$ LB (Lennox L) base in Milli-Q water and sterilised by autoclaving. Antibiotic(s) were added immediately prior to use.

Protein expression cultures

Cultures for protein expression were grown in 2 L to 3 L baffled conical flasks in a shaking incubator at 170 rpm to 180 rpm. Expression cultures were inoculated with an overnight culture $\approx \frac{1}{20}$ the volume and grown at 37 °C until induction of protein expression.

Transcription of the KDO8PS genes were under the control of an inducible *lac* operon, therefore protein expression was induced by addition of IPTG to a final concentration of either 0.5 or 1 mM, when cultures were at a mid-logarithmic phase (OD_{600} of 0.4 AU to 0.8 AU). After induction, cultures were kept shaking at 37 °C for 4 h before being harvested, or transferred to 23 °C or 25 °C and harvested the following morning.

Autoinduction

Autoinduction of protein expression was used in place of IPTG-based induction for some cultures. This technique was based on the methods of Studier.[93]

For an example 1 L culture, 10 g tryptone and 5 g yeast extract in 960 mL of water was autoclaved. Added to this was 20 mL of a 50× 5052 stock solution and of a 50× M stock solution, 1 mL of 1 M $MgSO_4$, 0.2 mL of a trace metal solution, and the appropriate antibiotic(s). The complete medium was then inoculated with ≈10 mL of an overnight pre-culture and incubated for 4 h at 37 °C before being transferred to 20 °C for 20 h to 36 h. The amounts and volumes were appropriately scaled depending on the culture volume.

50× 5052 stock: 25 % (w/v) glycerol, 2.5 % (w/v) glucose, and 10 % (w/v) lactose. Filter sterilised (0.2 µm).

50× M stock: 500 mL, 1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, 0.25 M Na_2SO_4. Autoclaved.

Trace metal solution: 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$/$ZnSO_4$, 2 mM $CoCl_2$, $CuCl_2$, $NiCl_2$, $NaMoO_4$, Na_2SeO_3 and H_3Bo_3 in ≈60 mM HCl.

Cell harvesting

The method used to harvest cells from cultures used for protein expression varied depending on rotor and centrifuge availability. Typically, large cultures were harvested in 0.8 L or 1 L bottles at 14 000 *g* for 15 min to 30 min at 4 °C. The pellets were then re-suspended in a small volume of the supernatant liquid and transferred to 50 mL centrifuge tubes and centrifuged at 12 000 *g* for 10 min at 4 °C. The supernatant liquid was removed and the cell pellet either immediately lysed, or stored at −80 °C.

Alternatively, cultures were sometimes progressively harvested in 50 mL centrifuge tubes at 12 000 *g* for 10 min at 4 °C. Or, in 0.4 L buckets at 4 618 *g* for

40 min at 4 °C, before being re-suspended in a small volume of the supernatant liquid and harvested in 50 mL centrifuge tubes as described above.

Small volume cultures (≤5 mL) were progressively harvested in 1.6 mL micro-centrifuge tubes at 16 000 *g* for 1 min.

Cell lysis

Cells were lysed by one of two methods: with a chemical detergent or by sonication. Sonication was performed with an Omni-Ruptor 4 000 Ultrasonic Homogenizer (Omni International). Cell pellets were resuspended in chilled lysis buffer (typically 10 mL) on ice, and sonicated in a beaker surrounded by packed ice by one of two sonication protocols. Sonicated at 70–80 percent power either for 6–8 multiples of 30 s with 1 min pauses between bursts, or by 3 multiples of 5 min with 30–40 percent pulsation.

Lysis by chemical detergent was performed using BugBuster® Protein Extraction Reagent (Novagen). Typically, cell pellets were resuspended in lysis buffer containing the BugBuster® reagent to a volume of 5 mL after which 4 µL of Benzonase® Nuclease (Novagen) (an engineered promiscuous endonuclease) was added and the solution shaken at 37 °C for at least 20 min.

The lysis buffer for *Nme*KDO8PS was 10 mM 1,3-bis[tris(hydroxymethyl)methylamino]propane (BTP) pH 7.5, 1 mM dithiothreitol (DTT), 200 mM KCl and 200 µM PEP. For *Afe*KDO8PS the lysis buffer was 10 mM BTP pH 7.2 and 1 mM EDTA.

Cellular debris were removed by centrifugation either at 16 000 *g*, 24 000 *g*, 30 000 *g* or 48 000 *g* for 30 min at 4 °C, depending on rotor and centrifuge availability.

Protein purification

All proteins were purified by a three-step procedure composed of, in step order, anion-exchange, hydrophobic-interaction and size-exclusion chromatography. The chromatography was performed on Bio-Rad Biologic DuoFlow and GE Healthcare AK̈TApurifier™ 10 machines.

Before use, all buffers were filtered (0.2 µm) under vacuum and degassed. Samples were all likewise filtered before being loaded to either a 10 mL or 50 mL Superloop™ (GE Healthcare), or, for samples of volume less than 5 mL, into a static loop. The eluate was fractionated in 2 mL fractions collected in 96-well plates. On AK̈TApurifier™ 10 machines, the eluate from injection of the sample onto the column (flow-through) was collected in 50 mL centrifuge tubes. Elution from the columns was monitored principally at 280 nm, and depending on the machine being used, also at 260 nm and 214 nm. When the fractions corresponding to KDO8PS could not be deduced from the chromatogram, suspected KDO8PS-containing fractions were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and/or for KDO8PS activity. Fractions containing KDO8PS were then pooled and prepared for the next chromatographic step.

Anion exchange chromatography

AEC was performed using SOURCE™ 15Q resin packed in a Tricorn 10/100 column (GE Healthcare), normally at 4 ℃. The supernatant liquid (resulting from removal of the cellular debris by centrifugation after cell lysis) was diluted four-fold with a no-salt buffer to reduce the concentration of anions. For *Nme*KDO8PS (pI is 6.25), the no-salt buffer was 10 mM BTP pH 7.5, and for *Afe*KDO8PS (pI is 5.54) was 10 mM BTP pH 7.2 and 1 mM EDTA. Elution from the column was performed using gradients of increasing NaCl concentration up to 1 M at a flow rate of 1 mL min^{-1} to 3 mL min^{-1}.

For *Nme*KDO8PS and mutant proteins thereof, the elution protocol was a 0 M to 0.1 M NaCl linear gradient over 25 mL following by 14 mL of

isocratic flow. Then the concentration of NaCl was raised in a linear gradient from 0.1 M to 0.2 M over 50 mL, followed by an isocratic flow for 20 mL, and then an isocratic flow at 1 M NaCl for 35 mL. *Nme*KDO8PS eluted between 0.16 M and 0.18 M NaCl.

For *Afe*KDO8PS and mutant proteins thereof, the elution protocol consisted of an isocratic flow at 0.14 M NaCl for 14 mL, followed by a linear gradient between 0.14 M and 0.32 M NaCl over 50 mL. An isocratic flow for 14 mL at 0.32 M and for 30 mL at 1 M NaCl followed. *Afe*KDO8PS eluted between 0.25 M and 0.28 M NaCl.

Hydrophobic interaction chromatography

The HIC was performed using SOURCE™ 15Phe resin packed also in a Tricorn 10/100 column (GE Healthcare). This purification step was performed at room temperature, as at 4 ℃ the hydrophobic interactions are weaker and the resin performance is degraded. Elution was performed at a flow rate of 1 mL min^{-1} to 3 mL min^{-1}.

Crystalline $(NH_4)_2SO_4$ was added to pooled AEC fractions to a concentration of 1 M before the sample was filtered and loaded into a loop. The same buffers as used in anion exchange were used for HIC, except $(NH_4)_2SO_4$ was substituted in place of NaCl. For all proteins, elution was by linear gradient from 1 M to 0 M $(NH_4)_2SO_4$ over 100 mL. *Nme*KDO8PS and mutants eluted, depending on flow rates and column condition, between 0.65 M and 0.35 M $(NH_4)_2SO_4$, whereas *Afe*KDO8PS and mutants eluted between 0.30 M and 0.14 M $(NH_4)_2SO_4$.

Size exclusion chromatography

Fractions containing KDO8PS from HIC were pooled and loaded into a loop. SEC was performed using a HiLoad 26/60 Superdex™ 200 prep grade column (GE Healthcare), at 4 ℃, in the same no-salt buffer used in the both AEC and HIC. The elution flow rate was 2.5 mL min^{-1}.

Polyacrylamide gel electrophoresis

SDS-PAGE was performed using either self-poured or pre-cast gels.

The preparation of self-poured SDS-PAGE gels was based on the method of Laemmli[115] and were composed of a 12 % (w/v) separating gel and a 4 % (w/v) stacking gel. Samples for analysis by SDS-PAGE (6 µL to 9 µL) were mixed with a protein loading buffer and sometimes boiled. Electrophoresis was performed using a Mini PROTEAN® 3 cell or Mini PROTEAN® Tetra Cell (Bio-Rad) at 200 V for 50 min using a Tris-glycine SDS running buffer.

Protein loading buffer (4×): 41.5 mM Tris-HCl pH 6.8, 40 % (v/v) glycerol, 8 % (w/v) sodium dodecyl sulfate (SDS), 10 mM EDTA, sufficient bromophenol blue to be visible and 10 µL of 0.5 M DTT added immediately before use.

Running buffer (5×): 1 M glycine, 0.25 M Tris, 1 % (w/v) SDS.

SDS-PAGE was also performed using NuPAGE® 10 % Bis-Tris Gel 1.0 mm × 12-well pre-cast gels (Invitrogen) in NuPAGE® MOPS SDS Running Buffer (Invitrogen). Electrophoresis was performed using a XCell SureLock™ Electrophoresis Cell (Invitrogen), also at 200 V for 50 min. Samples were mixed with NuPAGE® LDS Sample Buffer (4×) and a reducing agent, and again sometimes boiled. The reducing agent was either NuPAGE® Reducing Agent or DTT.

Visualisation

All SDS-PAGE gels were stained in square petri dishes, mostly by coomassie brilliant blue R-250, for at least 20 min before being destained to remove excess dye. Staining and destaining times were reduced by heating gel and solution in a microwave oven until almost boiling. Alternatively, some gels were stained using SimplyBlue™ Safe Stain (Invitrogen) instead of coomassie brilliant blue R-250. In the latter case, gels were first washed and also destained in water, and all steps were performed with heating until almost

boiling in a microwave oven.

Stain: 0.1 % (w/v) coomassie brilliant blue R-250, 10 % (v/v) acetic acid and 40 % (v/v) methanol

Destain: 10 % (v/v) acetic acid and 40 % (v/v) methanol

Molecular weight standards (Novex® Sharp Pre-Stained Protein Standards [Invitrogen] or SDS-PAGE Molecular Weight Standards, Low Range [Bio-Rad]) were run as the sample in one lane of each gel. Photographs of gels were taken on a white light transilluminator in a Molecular Imager® Gel Doc™ XR (Bio-Rad).

Concentrating and buffer exchanging of protein solutions

Protein solutions were concentrated using 10 000 Da molecular weight cut-off (MWCO) devices (Vivaspin 2 [GE Healthcare], Vivaspin 500 and Vivaspin 20 [Sartorius Stedim Biotech] and Amicon® Ultra-4 [Millipore]). All filtration units were rinsed before use with Milli-Q water. Protein solutions were buffer exchanged (e.g. into binding buffer for ITC) by repeatedly concentrating in a MWCO device and diluting with desired buffer.

Enzyme storage

Purified protein was divided into aliquots in either thin-walled PCR or 0.6 mL micro-centrifuge tubes in 30 µL to 200 µL aliquots, flash-frozen in liquid nitrogen, and stored at −80 °C. All protein samples were rapidly thawed immediately before use and kept on ice. Protein solutions were stored at 2 mg mL^{-1} to 40 mg mL^{-1}.

Protein concentration determination

For kinetic analyses and comparison to previous work, the concentration of purified protein solutions was measured by the method of Bradford.[116] Bio-Rad QuickStart Bradford Dye Reagent 1× (1 mL) was mixed with 20 µL of sample in a plastic cuvette and the absorption at 595 nm measured after 5 min using a water-blanked Cary 100 UV-visible spectrophotometer or Bio-Rad SmartSpec Plus spectrophotometer. A calibration curve, measured by the same method, created with known concentrations (0.0 mg mL^{-1}, 0.1 mg mL^{-1}, 0.2 mg mL^{-1}, 0.4 mg mL^{-1} and 0.8 mg mL^{-1}) of bovine serum albumin was used to convert absorption to concentration.

For all other purposes, the concentration was measured by absorption at 280 nm by an appropriately blanked Nanodrop ND-1 000 spectrophotometer, using 2 µL samples of purified protein solution. The concentration was calculated from the absorption using molar extinction coefficient values of 6 335 M^{-1} cm^{-1} and 17 335 M^{-1} cm^{-1} for *Nme*KDO8PS and *Afe*KDO8PS respectively. The coefficient values, calculated from protein sequences using the ProtParam tool on the ExPASy Proteomics Server,[117] were the same for both wild-type and mutant proteins.

Standard enzyme activity assay

The standard assay for KDO8PS and DAH7PS activity monitored the consumption of PEP by loss of absorbance at 232 nm, based on the method of Schoner and Herrmann.[118] All kinetic measurements were made using a Varian Cary 100 UV-visible spectrophotometer, stoppered 1 cm pathlength quartz cuvettes and a total assay volume of 1 mL. Standard assays were carried out at 37 ℃ and all assay solutions were equilibrated to temperature and baseline absorbance before initialisation with either enzyme, metal ion, or A5P. Standard reaction mixtures contained 50 mM BTP pH 7.2, a variable concentration of PEP, A5P and enzyme, and when appropriate 100 µM $MnSO_4$.

Assay buffers were treated with Chelex® 100 resin (Bio-Rad) prior to use to remove metal ions, and Chelex-treated buffer solution was used to create concentrated stock solutions of PEP and A5P. Metal-ion solutions were created using either Chelex-treated buffer or Milli-Q water.

Initial rates of reaction were measured as a least-squares fit of the initial rate data using Cary WinUV Kinetics Application (version 3.00, Varian). A unit of enzyme activity was defined as the loss of $1\,\mu mol\,min^{-1}$ of PEP at 37℃, and could be calculated from the measured loss of absorbance (at 232 nm) using Beer's Law (ϵ is $2.8\times10^3\,M^{-1}\,cm^{-1}$). The values of the kinetic parameters were determined by fitting data to the Michaelis-Menten equation using the software GraFit (version 5.0.13, Erithacus Software Limited). For cooperative kinetics, data were fitted to the Hill equation using Origin (version 8.1, OriginLab®).

Determination of substrate concentrations

The concentrations of substrate (e.g. PEP and A5P) solutions were determined using the standard assay system for KDO8PS activity. The concentration of the substrate not being measured was used in excess and the substrate for which the concentration was being measured was used at a limiting concentration. That the correct substrate was limiting was checked by addition of more of it after the reaction had reached completion. The change in absorbance, between that before initiation of the reaction and that at completion, was measured as ΔA_1. A control reaction, absent in one of the substrates, was used to measure the change in absorbance due to the addition of enzyme to the cuvette at initiation of the reaction (ΔA_2). The total change in absorbance ($\Delta A_1 + \Delta A_2$), was converted using Beer's Law to a concentration of the limiting substrate in the cuvette, and the stock substrate concentration calculated.

Measurement of protein mass

The masses of purified proteins were measured by electrospray ionisation using either a Micromass LCT Classic, Bruker maXis 3G, or a LTQ-Orbitrap XL (Otago University). Proteins were either buffer exchanged into a low concentration of ammonium bicarbonate, or were diluted from stored concentrations ($2\,\mathrm{mg\,mL^{-1}}$ to $40\,\mathrm{mg\,mL^{-1}}$) to $1\,\mathrm{mg\,mL^{-1}}$ with water.

Differential scanning fluorimetry

Melt temperatures of the proteins were determined in the presence of different additives by DSF, using an iCycler iQ5 Multicolour Real-Time PCR Detection System (Bio-Rad). The method used was based on that of Nordlund et al.[119] Triplicate protein samples were added with mixing to buffer (containing additives) and SYPRO orange dye (Invitrogen) in a 96-well microplate (*vide infra*), which was then sealed. The melt proceeded in 0.2 °C increments from 20 °C to 95 °C, with a 20 s dwell time after each temperature rise. Measurements of the fluorescence were made at the end of each dwell time.

For each sample (three replicates and blank) 100 µL of buffer (50 mM BTP pH 7.2) including any additive(s) was prepared, to which 5 µL of 250× SYPRO orange dye was added (condition solution). In the blank well for the sample, 21 µL of the condition solution was added to 4 µL of water or buffer. Then, 16 µL of protein (at $0.6\,\mathrm{mg\,mL^{-1}}$) was added to the remaining condition solution before 25 µL was dispensed into three replicate wells. In all experiments, each additive was at a final concentration of 80 µM.

For analysis an Excel spreadsheet with custom VBA-scripted macro was created to toggle and dynamically display melt and derivative curves for each sample. The melt temperatures were calculated as the temperature of maximum inflection of the melting curve after subtracting the reading of a blank well containing buffer and dye but lacking protein.

Circular dichroism

CD experiments were performed using a JASCO J-815 Spectropolarimeter. Unless otherwise described, all CD spectra were recorded from 260 nm to 190 nm using a 0.5 nm data pitch, 1 s response, 1 nm bandwidth at 25 °C. A 2 mL or 3 mL protein solution at a concentration of 0.03 mg mL^{-1} in a 3 mL quartz cuvette with 1 cm path length was used for each measurement. The buffer for all experiments was 10 mM phosphate pH 7.0. Blank spectra were recorded of buffer alone for each experiment before addition of protein.

The thermal denaturation was measured at 220 nm where a shift in the CD spectrum corresponding to α-helix content occurs. The CD was measured every 2 °C while heating from 25 °C to 90 °C in a stoppered cuvette, and the melt temperatures were calculated as the temperature of maximum inflection of the melting curve.

Data were smoothed using the Savitzky-Golay algorithm of the Jasco Spectra Manager™ (version 1.5) with a convolution width of 15.

Isothermal titration calorimetry

All ITC experiments were performed on a VP-ITC microcalorimeter (Microcal, GE Healthcare) operating at 298 K. All experiments consisted of 29 injections: one 2 µL injection and 28 subsequent 10 µL injections (unless otherwise stated), using a reference power of 10 µCal s^{-1}. For every experiment the ligand was in the syringe and protein in the cell. Before use, protein solutions were buffer exchanged into a binding buffer consisting either of 50 mM BTP pH 7.2 (*Nme*KDO8PS and mutants) or 50 mM BTP pH 7.2 and 0.5 mM $MnSO_4$ (*Afe*KDO8PS and mutants). The protein concentration was measured by UV absorption and all solutions were degassed in a vacuum immediately before use and both the cell and syringe were washed several times with binding buffer.

The initial datum point was routinely deleted to allow for diffusion of

ligand across the needle tip during the equilibration period. Heats of dilution experiments were measured independently and subtracted from the integrated data before curve fitting in Origin (version 7.0, OriginLab®) with the standard one-site model supplied by MicroCal.

Crystallisation of *Nme*KDO8PS and mutants

Crystals of mutant *Nme*KDO8PSs were grown by hanging-drop vapour diffusion in 24-well VDX plates (Hampton Research). Protein solutions [6 mg mL^{-1} to 30 mg mL^{-1}, in 10 mM BTP (pH 7.5)] were mixed 1:1 (v/v) with a reservoir solution containing 100 mM sodium acetate (pH 4.6; adjusted with glacial acetic acid) and 0.6 M to 3.0 M NaCl. The drop sizes were 2 µL, and the reservoir solution volume 500 µL.

The crystallisation trays were left at 20 °C until immediately before data collection, with crystals being transferred briefly into a cryoprotectant composed of 20 percent glycerol in the respective reservoir solution. Crystals typically began to form after 4 h and were fully formed in 24 h.

X-Ray data collection and structure determination

X-Ray diffraction data were collected at either The Institute of Fundamental Sciences at Massey University, or the Australian Synchrotron.

Protein was transported frozen to and plated in Palmerston North or Melbourne the day before collection. Pre-grown crystals were transported as hand luggage in foam-padding-adapted ice boxes containing the 24-well VDX plates. Alternatively, crystals were mounted on loops, flash-frozen in liquid nitrogen and placed in a sample storage cassette (Crystal Positioning Systems), which was stored in a dry shipper (Taylor-Wharton CX100) and transported as checked-in luggage.

At Massey University, a Rigaku MicroMax007 micro-focus copper rotating anode generator with AXCo PX70 focusing capillary optic (λ of 1.5418 Å)

coupled with an R-AxisIV^{++} image plate detector was used to collect data sets at 120 K (Oxford Cryosystems Series 700). Data collection and processing were performed with Crystal Clear™ (Rigaku). Data collection at the Australian Synchrotron used either the Macromolecular Crystallography (MX1) or Micro Crystallography (MX2) beamlines[120] and were processed using iMosflm and Scala (CCP4 suite).[121]

Structure determination and refinement

Structures of *Nme*KDO8PS mutant proteins were solved by molecular replacement, using either the structure of wild-type *Nme*KDO8PS (PDB code 2QKF) or *Nme*N59A (PDB code 3QPZ) (unless otherwise stated) as starting models (with all non-protein molecules removed) and performing three rounds of rigid-body refinement using Refmac5.[122] The same set of reflections for calculation of R_{free} was carried through.

Refinements were conducted with Refmac5,[122] and electron-density maps were analysed with Coot.[123] The validation tools of Coot and MolProbity[124] were used to check for, and correct, conformation infelicities.

9.2 Methods for Chapter 2

Gathering of sequences

Annotated DAH7PS and KDO8PS sequences were downloaded from BRENDA. Each set of sequences was initially aligned using ClustalX, and partial and other erroneously annotated sequences were removed. This resulted in a set of 2069 DAH7PS sequences and a set of 1005 KDO8PS sequences.

Categorisation of DAH7PS sequences into types

Each DAH7PS belongs to one of three distinct types (Iα, Iβ or II), where type Iβ are most similar to KDO8PS. It is known that Iβ DAH7PS can possess an N- or a C-terminal extension or no extension at all, while type Iα DAH7PS have both an N-terminal extension and an α5β6 loop extension. Type II DAH7PS always have an N-terminal extension, a KPRS motif, and often an α2β3 loop extension, though this is sometimes absent.

From MSAs of test sets of DAH7PSs of known types, additional sequence features were identified that could be used as type-fingerprints. Unique to type I sequences were a G(A/S)R(N/T/S) (GARN) motif, and two short motifs K(R/N) and RG. The α2β3 loop is bound by a KPR(S/T) motif and the GARN motif, while the α5β6 loop is bound by K(R/N) and RG.

To identify the Iβ DAH7PS sequences, a Ruby script that implemented rule-based scoring was developed to assign one of the three types (Iβ, Iα or II) to each sequence based on the fingerprints identified above (Table 9.1). Briefly, an index position in each motif was assigned by finding the location of the KPR(S/T) motif. If the GARN motif was found, then the length of the α2β3 loop was measured as the distance between KPR(S/T) and this motif. If the GARN motif was present, the K(R/N) and RG motifs were located in a search range indexed from the GARN motif sequence position and the distance measured as the α5β6 loop length. The presence of KPRS rather

Table 9.1: Sequence-feature scoring system

Rule	Type	Points
KPRS rather than KPRT	II	1
No GARN motif	II	2
	I	−1
GARN motif present	I	1
	II	−1
α2β3 length		
length < 57	Iβ	2
57 < length < 70	Iα	2
α5β6 length		
12 < length < 28	Iβ	2
30 < length < 46	Iα	2

than KPRT, and the absence of the GARN motif increased the probability of a type II sequence.

Of the DAH7PS sequences analysed, 375 were assigned as type Iβ and these were independently aligned using ClustalX. Although on test-set data the sequences were unambiguously and correctly identified, the sequences assigned as Iβ were checked manually in the MSA by comparing the N-terminal extensions. Sequences could be grouped into three distinct types based on length and conserved residues, consistent with the known possibilities of the extension being absent, a chorismate mutase, or an ACT domain. The analysis between KDO8PS and Iβ DAH7PS compares only the barrel domain portion of the sequences, which means the various extensions did not need to be discarded.

Profile hidden Markov models

To create a profile HMM of each set of aligned sequences, the hmmbuild program of the HMMER3 package was used. The output of this program is an HMMER3 profile HMM file, which consists of position-specific scores for

each possible amino acid, as well as scores for the likelihood of any possible amino acid being inserted at each position. The latter represents gaps in the original MSA.

Most documentation and other programs that make use of profile HMM data assume the HMMER2 profile HMM file format. To make the data easier to process, the profiles were converted back to a close approximation of HMMER2 models and into the HMMER2 profile HMM file format, using the HMMER3 program hmmconvert.

HMM-HMM pairwise alignments were created using the hhalign program from the HHsearch package. The output is a query profile HMM aligned to a template profile HMM in a FASTA-like format. For both the query and template profile HMM, the output includes the core probability model sequence and the consensus sequence, where upper-case letters represent well conserved positions and lower case partially conserved positions. Approximately, amino acids with $\geq$60 percent probability are written as capital letters, and those with $\geq$40 percent probability with lower-case letters.

Converting position-specific scores to relative entropy

The HMMER2 profile HMM file format holds a score for each amino acid at each sequence position (a "match emission"). In addition it contains a null model, which consists of a "null emission" score (generic to all positions) for each amino acid. To compare the sequence conservation in each profile HMM, the position-specific scores for each amino acid were converted to relative entropy.[i] Each score (b_j) is stored as an 'entropy' or 'information':

$$b_j = 1000 \times log_2 \left(\frac{p_j}{q_j} \right) \tag{9.1}$$

where p_j is the probability of amino acid j being at a certain position, and q_j is the probability of an amino acid at the same position from the background

[i] The insertion emission was considered unimportant for this analysis.

distribution (the null model, which in the case of null emissions $q_j = \frac{1}{20}$). The relative entropy at each position (RE) was calculated as:

$$RE = \sum_{j=1}^{20} p_j b_j = \sum_{j=1}^{20} p_j 1000 \times log_2 \left(\frac{p_j}{q_j} \right) \tag{9.2}$$

and RE was plotted against position.

9.3 Methods for Chapter 3

Analytical ultracentrifugation

The analytical ultracentrifugation (AUC) experiments were performed using absorbance optics in a Beckman Coulter XL-I analytical ultracentrifuge equipped with UV/Vis scanning optics at 20 °C in 10 mM BTP pH 7.5 and 50 mM NaCl. The experiments were performed at 40 000 *g* and the wavelength for wild-type *Afe*KDO8PS was 237 nm and for *Afe*C21N was 236 nm. Radial scans were collected between 5.8 cm and 7.2 cm with 0.002 cm increments. The data were analysed according to a continuous size distribution model using SEDFIT. The resolution was 100 species, between 0.5 S and 15 S, $P = 0.95$, $\bar{v} = 0.7332$ mL g^{-1}, $\rho = 1.0001$ g mL^{-1} and $\eta = 1.0093$ mPa s^{-1}, calculated using SEDNTERP. Protein samples were 380 µL at 1 mg mL^{-1} and the reference cell volume was 400 µL. 12 mm double sector cells with quartz windows were used.

Kinetics

The assays for determining the wild-type *Afe*KDO8PS kinetic parameters included 100 µM Mn^{2+} or Cd^{2+} and 12 µg of enzyme. To determine K_m^{PEP}, the A5P concentration was fixed at 200 µM while the PEP concentration varied. For determination of K_m^{A5P}, the PEP concentration was fixed at 150 µM while the A5P concentration was varied.

The value of K_m^{2dR5P} was determined with 100 µM Mn^{2+}, 100 µM PEP and varying concentrations of 2dR5P, using 47 µg of wild-type *Afe*KDO8PS. Activity of wild-type *Afe*KDO8PS with glucose 6-phosphate (G6P) and E4P used 100 µM of PEP and 3 mM G6P and 1 mM E4P.

Activity of *Afe*C21N was measured using 100 µM PEP and 1 mM A5P, with and without $MnSO_4$. Activity with 2dR5P and R5P was measured used 1 mM and 3 mM respectively of each aldose substrate, and 100 µM of PEP.

Metal activation

The metal-activation profile of wild-type *Afe*KDO8PS was determined by standard enzyme assays containing 100 µM PEP, and 100 µM A5P, and 100 µM of $MnSO_4 \cdot 4\,H_2O$, $CoCl_2 \cdot 6\,H_2O$, $2\,CdCl_2 \cdot 5\,H_2O$, $FeSO_4 \cdot 7\,H_2O$, $ZnCl_2$, $CaCl_2 \cdot 2\,H_2O$, $BaCl_2 \cdot 2\,H_2O$, $SrCl_2 \cdot 6\,H_2O$, $NiCl_2 \cdot 6\,H_2O$, $MgSO_4 \cdot 7\,H_2O$ or $CuSO_4 \cdot 5\,H_2O$ in 50 mM BTP buffer containing 10 µM EDTA, pH 7.2 at 37 ℃. The enzyme (9.3 µg) was incubated with the metal ion for 5 min prior to initiation of the reaction by A5P.

Isothermal titration calorimetry

For the titration of PEP with *Nme*KDO8PS, 700 µM PEP in the syringe was injected into a solution of 50 µM *Nme*KDO8PS. For PEP with *Afe*KDO8PS, 720 µM PEP was injected into a solution of 45 µM *Afe*KDO8PS. For the titration of A5P, the concentrations of A5P and protein used were 0.65 mM A5P and 60 µM *Nme*KDO8PS. For *Afe*KDO8PS, 4 mM A5P was titrated into 0.2 mM *Afe*KDO8PS.

Titration of R5P with *Nme*KDO8PS (45 µM) in the presence of PEP (0.5 mM) used 0.7 mM R5P. For *Afe*KDO8PS, 4 mM R5P was titrated into 170 µM *Afe*KDO8PS in the presence of 0.5 mM PEP.

For titration of PEP with *Afe*C21N, the binding buffer was 50 mM BTP pH 7, the syringe concentration of PEP was 650 µM, and the protein concentration was 60 µM. The titration of A5P with *Afe*C21N comprised 23 injections, one 2 µL injection followed by 22 10 µL injections, of 1 mM A5P into a solution of 60 µM protein.

Crystallisation

Crystallisation conditions were screened for *Afe*KDO8PS using the PACT premier™ (Molecular Dimensions) and JCSG-plus™ (Molecular Dimensions)

Table 9.2: Crystallisation screens used at the C3

Protein	Plate	Screen	Additives	Temp. (℃)
*Afe*KDO8PS	MA002255	C3-2	10 µM $CdCl_2$, 0.2 mM PEP	8
	MA002256	C3-3	10 µM $CdCl_2$, 0.2 mM PEP	8
	MA002257	C3-4	10 µM $CdCl_2$, 0.2 mM PEP	8
	MA002258	PACT Suite	10 µM $CdCl_2$, 0.2 mM PEP	8
	MA002260	RAND[a]	10 µM $CdCl_2$, 0.2 mM PEP	8
	MC003951	JCSG-plus	–	8
	MC003952	PACT Suite	–	8
*Nme*KARS	MC004431	C3-2	2 mM PEP	20
	MC004432	C3-3	2 mM PEP	20
	MC004433	C3-4	2 mM PEP	20
	MC004434	PACT Suite	2 mM PEP	20
	MC004435	C3-2	2 mM PEP	8
	MC004436	C3-3	2 mM PEP	8
	MC004437	C3-4	2 mM PEP	8

[a] Random screen around PEG-based conditions generated by C3.

screens, 24-well VDX-plates and hanging-drop vapour diffusion. The reservoir volume was 500 µL and the drop size was 2 µL (equal volume protein solution and reservoir). Crystallisation plates were incubated at 20 ℃.

*Afe*KDO8PS was also screened in 96-well plates employing sitting-drop vapour diffusion with Structure Screen I and II (Molecular Dimensions). This screening was performed in the absence of added metal ion, and with $MnSO_4$, with Cd-acetate, using 10 mg mL^{-1} protein. The drop size was 1 µL (1:1) and the reservoir volume 40 µL.

Lastly, *Afe*KDO8PS was also screened using the facilities of the C3 at CSIRO in Melbourne. Drop sizes were 300 or 350 pL (1:1). Details of the screening performed are listed in Table 9.2. The C3 was also used to attempt to find new crystallisation conditions for *Nme*KDO8PS.

Crystals which grew in the screens were tested for diffraction at either Massey University or the Australian Synchrotron. Conditions in which

promising crystals grew (Table 9.3) were screened around using hanging-drop vapour diffusion in 24-well plates. The optimisation varied protein concentration, component concentration (invariably PEG and salt), and trialled the effect of the presence of PEP and metal ion (none, Mn^{2+} or Cd^{2+}).

Table 9.3: Promising crystallisation conditions for *Afe*KDO8PS.

Screen	Location	Components
PACT	D9	0.2 M $LiCl_2$, 0.1 M Tris pH 8.0, 20 % (w/v) PEG 6 000
	E1	0.2 M NaF, 20 % (w/v) PEG 3 350
	E2	0.2 M NaBr, 20 % (w/v) PEG 3 350
	E4	0.2 M KSCN, 20 % (w/v) PEG 3 350
	E9	0.2 M K-tartrate, 0.2 M Na-tartrate, 20 % (w/v) PEG 3 350
	E12	0.2 M Na-malonate, 20 % (w/v) PEG 3 350
	G6	0.2 M Na-formate, 20 % (w/v) PEG 3 350, 0.1 M BTP pH 7.5
	G7	0.2 M Na-acetate, 20 % (w/v) PEG 3 350, 0.1 M BTP pH 7.5
	H12	0.2 M Na-malonate, 20 % (w/v) PEG 3 350, 0.1 M BTP pH 8.5
JCSG+	A12	0.2 M KNO_3, 20 % (w/v) PEG 3 350
	B2	0.2 M NaSCN, 20 % (w/v) PEG 3 350
	E12	0.1 M imidazole pH 8.0, 10 % (w/v) PEG 8 000
	G1	0.1 M HEPES pH 7.0, 30 % (w/v) Jeffamine ED-2001
	G6	0.2 M Na-malonate pH 7.0, 20 % (w/v) PEG 3 350
C3-2	A7	0.05 M Tris pH 8.0, 10 % (w/v) PEG 8 000, 10 % (w/v) PEG 1 000, 0.8 M Na-formate
	B11	0.1 M CHES pH 9.5, 25 % (w/v) PEG 1 500, 0.2 M $MgSO_4$
	E8	0.1 M Na-cacodylate pH 6.5, 10 % (w/v) PEG 3 000, 0.2 M $MgCl_2$
C3-3	B9	0.2 M Na_2-tartrate, 20 % (w/v) PEG 3 350
	G2	0.8 M Na-formate, 15 % (w/v) PEG 4 000, 0.05 M Tris pH 8.0
Hampton Crystal Screen I	C1	0.2 M tri-sodium citrate dihydrate, 0.1 M Tris-HCl pH 8.5, 30 % (w/v) PEG 400
	C4	0.1 M Na-HEPES pH 7.5, 1.5 M $LiSO_4$ monohydrate
	C5	0.2 M $LiSO_4$ monohydrate, 0.1 M Tris-HCl pH 8.5, 30 % (w/v) PEG 4 000
	H2	0.2 M Mg-formate

9.4 Methods for Chapter 4

Site-directed mutagenesis

The *Afe*KARS (*Afe*N57del) and *Nme*KARS (*Nme*N59del) constructs were created by site-directed mutagenesis that deleted the codon for the Asn residue of the KANRS motif using wild-type *Nme*KDO8PS plasmid as the template. To create the KPRS variants of both enzymes, the KARS construct was used as the template, and the Ala was mutated to Pro.

The mutagenesis to create the *Afe*K55A, *Afe*N57A, *Afe*N57D, *Nme*K57A, *Nme*N59A, and *Nme*N59D constructs had previously been performed by Jeffrey Yeoman (for those of *Afe*KDO8PS[84]) and Dr Fiona Cochrane (for those of *Nme*KDO8PS). The primers used to create these constructs (using the wild-type sequences as the template) are also listed below for posterity.

For the primer sequences below, the bases are split into codons and the mismatched bases are in lower-case letters.

*Afe*K55A

For: 5' `GT TCC TAC GAT gcA GCG AAC CGT TCT TCG`

Rev: 5' `CGA AGA ACG GTT CGC Tgc ATC GTA GGA AC`

*Afe*N57A

For: 5' `CC TAC GAT AAA GCG gca CGT TCT TCG GGG CAG`

Rev: 5' `CTG CCC CGA AGA ACG tgc CGC TTT ATC GTA GG`

*Afe*N57D

For: 5' `C TAC GAT AAA GCG gAC CGT TCT TCG GG`

Rev: 5' `CC CGA AGA ACG GTc CGC TTT ATC GTA G`

*Afe*N57del (KARS)

For: 5' `C TAC GAT AAA GCG *** CGT TCT TCG GGGC`

Rev: 5' `G CCC CGA AGA ACG *** CGC TTT ATC GTA G`

*Afe*A56P (KPRS)

For: 5' `GT TCC TAC GAT AAA cCG CGT TCT TCG GGG C`

Rev: 5' G CCC CGA AGA ACG CGg TTT ATC GTA GGA AC

*Nme*K57A

For: 5' C TTT AAA GCC TCT TTC GAC gcg GCA AAC CGT TC

Rev: 5' GA ACG GTT TGC cgc GTC GAA AGA GGC TTT AAA G

*Nme*N59A

For: 5' C TCT TTC GAC AAG GCA gca CGT TCC TCC ATC CAT TC

Rev: 5' GA ATG GAT GGA GGA ACG tgc TGC CTT GTC GAA AGA G

*Nme*N59D

For: 5' CT TTC GAC AAG GCA gac CGT TCC TCC ATC

Rev: 5' GAT GGA GGA ACG gtc TGC CTT GTC GAA AG

*Nme*N59del (KARS)

For: 5' CT TTC GAC AAG GCA *** CGT TCC TCC ATC C

Rev: 5' G GAT GGA GGA ACG *** TGC CTT GTC GAA AG

*Nme*A58P (KPRS)

For: 5' CC TCT TTC GAC AAG ccg CGT TCC TCC ATC C

Rev: 5' G GAT GGA GGA ACG cgg CTT GTC GAA AGA GG

Kinetics

To determine the kinetic parameters for each enzyme, the following concentrations of substrates were used. For *Nme*N59A $K_{\mathrm{m}}^{\mathrm{A5P}}$ 100 µM PEP and 88.5 µM to 3540 µM A5P, and for $K_{\mathrm{m}}^{\mathrm{PEP}}$ 788 µM A5P and 4.97 µM to 99.4 µM PEP. For *Nme*KARS $K_{\mathrm{m}}^{\mathrm{A5P}}$ 100 µM PEP and 210 µM to 7965 µM A5P, and for $K_{\mathrm{m}}^{\mathrm{PEP}}$ 800 µM A5P and 10 µM to 100 µM PEP. For *Nme*KPRS $K_{\mathrm{m}}^{\mathrm{A5P}}$ 115 µM PEP and 15.75 µM to 525 µM A5P, and for $K_{\mathrm{m}}^{\mathrm{PEP}}$ 158 µM A5P and 4.97 µM to 99.4 µM PEP. For *Afe*N57A $K_{\mathrm{m}}^{\mathrm{A5P}}$ 100 µM PEP and 27.6 µM to 345 µM A5P, and for $K_{\mathrm{m}}^{\mathrm{PEP}}$ 800 µM A5P and 1.95 µM to 26.0 µM PEP. Substrate concentrations of up to 200 µM PEP and 1 mM A5P were used to test for activity of those mutants with no perceptible activity.

Substrate specificity

Substrate specificity was assessed using the standard assay, using 200 µM PEP, and up to 1 mM alternate aldose substrate. The substrate concentrations for determining the kinetic parameters of 2dR5P with *Nme*N59A were 100 µM PEP and 86.6 µM to 1732 µM 2dR5P and for *Afe*N57A 100 µM PEP and 78 µM to 3900 µM 2dR5P.

Isothermal titration calorimetry

For *Nme*K57A, the syringe concentration of PEP was 650 µM, and the concentration of protein in the cell was 50 µM. For *Afe*K55A the PEP concentration in the syringe was 1.5 mM, and the protein concentration in the cell was 150 µM.

Crystallisation, structure determination and refinement

The X-ray diffraction data for *Nme*K57A, *Nme*N59A, and *Nme*KARS were collected at Massey University, and those for *Nme*KPRS at the Australian Synchrotron on the Micro Crystallography (MX2) beamline.

Data were collected from crystals of *Nme*K57A grown in 2.2 M NaCl, *Nme*N59A grown in 1.0 M NaCl, *Nme*KARS grown in 1.4 M NaCl, and *Nme*KPRS grown in 2.6 M NaCl. All four mutant proteins crystallise like the wild-type protein in orthorhombic space group $P2_12_12_1$ and diffracted to 1.95 Å, 1.75 Å, 1.90 Å, and 2.70 Å, respectively, with the following unit cell dimensions: $a \approx 82$ Å, $b \approx 85$ Å, and $c \approx 163$ Å. The results are summarised in Tables 4.4 and 4.5, along with key structure refinement details.

9.5 Methods for Chapter 5

Site-directed mutagenesis

The constructs for three of the single mutant proteins (*Afe*S207A, *Nme*Q202A and *Nme*S211A) were created by Ben Gloyne and Dr Richard Hutton using the primers listed below.[95] *Afe*Q198A was created using wild-type *Afe*KDO8PS was the template. The same primers were used to create the double-mutant proteins, for *Afe*Q198A/S207A using *Afe*Q198A as the template plasmid, and for *Nme*Q202A/S211A using *Nme*S211A as the template plasmid, and the respective primers to add the other mutation.

The β7α7-truncated constructs were created by Evan Nimmo and Dr Richard Hutton.[95,96] To create the L7trun/KPRS constructs the β7α7-truncated plasmids were used as template on which the Asn of the KANRS was deleted. The product of this mutagenesis (KARS) was then used as the template for subsequent mutagenesis which mutated the Ala to Pro (creating KPRS), using the same protocols as described in Chapter 4.

For the primer sequences below, the bases are split into codons and the mismatched bases are in lower-case letters.

*Afe*Q198A

For: 5' `GAT GCG ACC CAT TCC GTA gcG CTC CCC GGT G`

Rev: 5' `C ACC GGG GAG Cgc TAC GGA ATG GGT CGC ATC`

*Afe*S207A

For: 5' `AG GGT GAC CGG gCC GGA GGC CAG`

Rev: 5' `CTG GCC TCC GGc CCG GTC ACC CT`

*Nme*Q202A

For: 5' `GTT ACC CAT TCC CTG gcA ACC CGC GAT GCC GG`

Rev: 5' `CC GGC ATC GCG GGT Tgc CAG GGA ATG GGT AAC`

*Nme*S211A

For: 5' `GCC GGT TCT GCC GCA gCC GGC GGT`

Rev: 5' `ACC GCC GGc TGC GGC AGA ACC GGC`

Kinetics

The assays for determining the kinetic parameters kept the concentration of one substrate constant while varying that of the other and vice versa. The amount of KDO8PS used in each assay ranged from 0.02 mg to 0.08 mg depending on the experiment. To determine $K_{\mathrm{m}}^{\mathrm{PEP}}$, the A5P concentration was kept at 150 µM, and the concentration of PEP was varied from 2.8 µM to 98 µM for all Q to A and S to A mutants, except for *Afe*Q198A/S207A and *Afe*S207A for which the A5P concentration was 854 µM and the PEP concentration varied from 8.5 µM to 204 µM and from 4.3 µM to 85 µM, respectively. For *Nme*Q202A/S211A, the A5P concentration was 1768 µM and the PEP concentration varied between 2.6 µM and 63 µM.

To determine $K_{\mathrm{m}}^{\mathrm{A5P}}$ for these mutants, the PEP concentration was kept at 100 µM and the A5P concentration was varied between 6.4 µM and 205 µM, 102 µM and 411 µM, and 26 µM and 411 µM for *Nme*S211A, *Nme*Q202A, and *Afe*Q198A, respectively. For *Afe*Q198A/S207A, *Afe*S207A, and *Nme*Q202A/S211A, the PEP concentration was 150 µM and the A5P concentration was varied between 24 µM and 488 µM for both *Afe*Q198A/S207A and *Afe*S207A and between 110 µM and 384 µM for *Nme*Q202A/S211A.

For the $K_{\mathrm{m}}^{\mathrm{A5P}}$ determinations for the L7trun enzymes, the PEP concentration was kept constant at 150 µM and the A5P concentration was varied between 500 µM and 3 mM for *Nme*L7trun and between 200 µM and 3 mM for *Afe*L7trun. The amount of KDO8P produced by the L7trun proteins was measured using the periodate-thiobarbituric acid assay[15] to rule out uncoupled hydrolysis of PEP. The two L7trun mutant proteins produced the same amount of KDO8P as the wild-type enzymes. Assays for determining the activity of the KPRS/L7trun enzymes used 150 µM PEP and 1 mM A5P.

Isothermal titration calorimetry

For *Afe*L7trun, the protein concentration in the cell was 65 µM, and the PEP concentration in the syringe was 650 µM. For *Nme*L7trun, the protein

concentration in the cell was 110 µM, and the PEP concentration in the syringe was 2 mM.

Crystallisation, structure determination and refinement

The X-ray diffraction data for *Nme*Q202A and *Nme*S211A were collected at Massey University, and those for *Nme*L7trun and *Nme*KPRS/L7trun at the Australian Synchrotron on the Macromolecular Crystallography (MX1) beamline.

Data were collected for crystals of *Nme*Q202A grown in a 0.8 M NaCl condition, *Nme*S211A in 1.0 M NaCl, *Nme*L7trun in 1.6 M NaCl, and *Nme*-KPRS/L7trun in 2.8 M NaCl. All four mutant proteins crystallise like the wild-type *Nme*KDO8PS in orthorhombic space group $P2_12_12_1$ and diffracted to 2.05 Å, 1.90 Å, 1.90 Å, and 2.20 Å, respectively, with the following unit cell dimensions for three of the four: $a \approx 82$ Å, $b \approx 85$ Å, and $c \approx 163$ Å. The exception was *Nme*KPRS/L7trun, which had the following unit cell dimensions: $a = 82$ Å, $b = 104$ Å, and $c = 150$ Å.

A modified wild-type structure with the excised β7α7 loop residues removed from the PDB file was used for the solution of *Nme*L7trun, carrying through the same set of reflections for calculation of R_{free}. The *Nme*L7trun model was used with Phaser[125] to determine the structure of *Nme*KPRS/L7trun by molecular replacement, with the same R_{free} reflection set. The results are summarised in Tables 5.3 and 5.4, along with key structure refinement details.

9.6 Methods for Chapter 6

Site-directed mutagenesis

The constructs for the mutant proteins in this chapter were created by site-directed mutagenesis using wild-type construct plasmids as the templates. For the primer sequences below, the bases are split into codons and the mismatched bases are in lower-case letters.

*Afe*D90A

For: 5' G GTG CCG GTC GTT ACG Gcg GTT CAT GAA AAA GAG G

Rev: 5' C CTC TTT TTC ATG AAC cgC CGT AAC GAC CGG CAC C

*Afe*D90N

For: 5' GG GTG CCG GTC GTT ACG aAT GTT CAT GAA AAA GAG

Rev: 5' CTC TTT TTC ATG AAC ATt CGT AAC GAC CGG CAC CC

*Afe*D90E

For: 5' GTG CCG GTC GTT ACG GAa GTT CAT GAA AAA GAG G

Rev: 5' C CTC TTT TTC ATG AAC tTC CGT AAC GAC CGG CAC

*Nme*D92A

For: 5' C ATC CCC GTC ATT ACC Gcg GTA CAC GAA CCC CAT C

Rev: 5' G ATG GGG TTC GTG TAC cgC GGT AAT GAC GGG GAT G

*Nme*D92N

For: 5' GGC ATC CCC GTC ATT ACC aAT GTA CAC GAA CCC CAT CAG

Rev: 5' CTG ATG GGG TTC GTG TAC ATt GGT AAT GAC GGG GAT GCC

*Nme*D92E

For: 5' C CCC GTC ATT ACC GAa GTA CAC GAA CCC CAT C

Rev: 5' G ATG GGG TTC GTG TAC tTC GGT AAT GAC GGG G

Kinetics

To determine $K_{\mathrm{m}}^{\mathrm{PEP}}$ for *Nme*D92A, the A5P concentration was 0.5 mM and the concentration of PEP was varied between 11 µM and 222 µM. To determine $K_{\mathrm{m}}^{\mathrm{A5P}}$ for *Nme*D92A, the PEP concentration was 150 µM and the concentration of A5P was varied between 158 µM and 6320 µM. To determine $K_{\mathrm{m}}^{\mathrm{PEP}}$ for *Nme*D92N the concentration of A5P was 500 µM and PEP varied between 11 µM and 222 µM. To determine $K_{\mathrm{m}}^{\mathrm{A5P}}$, PEP was 150 µM and A5P varied between 7.9 µM and 1580 µM. To determine $K_{\mathrm{m}}^{\mathrm{PEP}}$ for *Nme*D92E, the A5P concentration was 5 mM and the concentration of PEP varied between 9.7 µM and 387 µM. To determine $K_{\mathrm{m}}^{\mathrm{A5P}}$, PEP was 340 µM and A5P varied between 218 µM and 7000 µM.

Assays to determine activity of *Afe*D90A and *Afe*D90N used 200 µM PEP and 500 µM A5P. To determine $K_{\mathrm{m}}^{\mathrm{PEP}}$ for *Afe*D90E, the concentration of A5P was 6 mM and the concentration of PEP was varied between 48 µM and 1452 µM. To determine $K_{\mathrm{m}}^{\mathrm{A5P}}$, PEP was 500 µM and A5P varied between 131 µM and 3722 µM.

Isothermal titration calorimetry

For each mutant protein the concentration of PEP in the syringe was 1.5 mM. The concentrations of protein in the cell were: *Afe*D90A, 15.7 µM; *Afe*D90N, 30.5 µM; *Afe*D90E, 21 µM; *Nme*D92A, 26 µM; *Nme*D92N, 40 µM; *Nme*D92E, 40.7 µM.

Crystallisation, structure determination and refinement

The X-ray diffraction data for *Nme*D92A, *Nme*D92N and *Nme*D92E were collected at the Australian Synchrotron on the Micro Crystallography (MX2) beamline. Data were collected for crystals of *Nme*D92A grown in a 1.0 M

NaCl condition, *Nme*D92N in 1.0 M NaCl, and *Nme*D92E in 1.4 M NaCl. All three proteins crystallise like wild-type *Nme*KDO8PS in orthorhombic space group $P2_12_12_1$, with the following unit cell dimensions: $a \approx 82$ Å, $b \approx 85$ Å, and $c \approx 163$ Å. The results are summarised in Tables 6.3 and 6.4, along with key structure refinement details.

9.7 Methods for Chapter 7

Site-directed mutagenesis

Site-directed mutagenesis was used to create the constructs of the mutant proteins characterised in this chapter. The template for the creation of *Nme*F114R/R117A and *Nme*F114R/R117Q was *Nme*F114R plasmid. To create *Nme*F114R/R117Q/F139G, *Nme*F114R/R117Q was used as the plasmid template. The primers that Dr Fiona Cochrane used to create *Nme*F114R, *Nme*F114A, *Nme*R117K, and *Nme*R117A are listed below for posterity.

*Nme*Quin was constructed from the *Nme*KPRS/L7trun template, first by addition of F139G, then F114R, and finally R117Q. The primers to add F114R used a different codon for Arg than what the primers for F114R/R117Q expected. This meant that the addition of R117Q contained an extra, unplanned base change.

Only one sequenced colony correctly incorporated the final desired mutation, however this construct also contained the undesired base change G376A, which translates into the mutation A126T. Further mutagenesis was attempted to correct this unwanted change in sequence (Fixup primers), but these attempts were unsuccessful.

For the primer sequences below, the bases are split into codons and the mismatched bases are in lower-case letters.

*Nme*R117Q

For: 5' `CC GCC TTT CTT GCG Cag CAG ACC GAT TTA G`

Rev: 5' `C TAA ATC GGT CTG ctG CGC AAG AAA GGC GG`

*Nme*F139G

For: 5' `C AAC ATC AAA AAA CCT CAG ggC CTC AGC CCC TCT CAA ATG`

Rev: 5' `CAT TTG AGA GGG GCT GAG Gcc CTG AGG TTT TTT GAT GTT G`

*Nme*F114R/R117A

For: 5' `CGC CTT GCG gca CAG ACC GAT TTA GTG G`

Rev: 5' C CAC TAA ATC GGT CTG tgc CGC AAG GCG

*Nme*F114R/R117Q

For: 5' GCC CGC CTT GCG CaG CAG ACC GAT TTA G

Rev: 5' C TAA ATC GGT CtG CTG CGC AAG GCG GGC

*Nme*F114R

For: 5' TGC GAT GTC ATC CAG CTT CCC GCC cgt CTT GC

Rev: 5' GC AAG acg GGC GGG AAG CTG GAT GAC ATC GCA

*Nme*F114R (Dr Fiona Cochrane)

For: 5' TGC GAT GTC ATC CAG CTT CCC GCC cgc CTT GC

Rev: 5' GC AAG gcg GGC GGG AAG CTG GAT GAC ATC GCA

*Nme*F114A

For: 5' TGC GAT GTC ATC CAG CTT CCC GCC gcg CTT GC

Rev: 5' GC AAG cgc GGC GGG AAG CTG GAT GAC ATC GCA

*Nme*R117A

For: 5' C GCC TTT CTT GCG gct CAG ACC GAT TTA GTG

Rev: 5' CAC TAA ATC GGT CTG agc CGC AAG AAA GGC G

*Nme*R117K

For: 5' CCC GCC TTT CTT GCG aaG CAG ACC GAT TTA GTG

Rev: 5' CAC TAA ATC GGT CTG Caa CGC AAG AAA GGC GGG

*Nme*Fixup

For: 5' GAT TTA GTG GTT GCC ATG gCA AAA ACT GGC AAC GTC

Rev: 5' GAC GTT GCC AGT TTT TGc CAT GGC AAC CAC TAA ATC

Protein purification

The mutant proteins of *Nme*KDO8PS were purified according to the purification procedures for wild-type *Nme*KDO8PS.[39] For the purification of *Nme*Quin, 50 mM NaCl and 0.5 mM PEP was added to all purification buffers.

Kinetics

To determine K_m^{A5P} for *Nme*R117Q, the concentration of A5P was varied between 111 µM and 5560 µM while the PEP concentration remained constant at 100 µM. To determine K_m^{PEP} for *Nme*F139G, PEP was varied between 7.8 µM and 156 µM with A5P at 1 mM, and for K_m^{A5P}, the concentration of A5P was varied between 39.5 µM and 1580 µM while PEP was 150 µM. To determine K_m^{A5P} for *Nme*F114R/R117A the concentration of A5P was varied between 133 µM and 3330 µM while PEP remained constant at 100 µM, and for *Nme*F114R/R117Q the concentration of A5P was varied between 172 µM and 2580 µM while PEP was also 100 µM. To determine K_m^{A5P} for *Nme*F114R/R117Q/F139G the A5P concentration was varied between 133 µM and 2931 µM while PEP was held at 100 µM.

Assays to determine the activity of *Nme*Quin were performed at 25 °C and used 200 µM of PEP and 1 mM of 2dR5P, R5P or E4P. Assays to determine activity with A5P used a concentration of 1 mM A5P and 160 µM PEP. Assays to determine the activity of *Nme*F114R/R117A, *Nme*F114R/R117Q, and *Nme*F114R/R117Q/F139G with 2dR5P, R5P and E4P used 1 mM of the aldose substrate, and 100 µM of PEP.

Isothermal titration calorimetry

The concentration of *Nme*Quin in the cell was 30 µM and the concentration of PEP in the syringe was 1.5 mM.

Crystallisation, structure determination and refinement

The X-ray diffraction data for *Nme*F139G, *Nme*F114A, *Nme*F114R, and *Nme*R117K were collected at Massey University, and those for *Nme*F114R, *Nme*R117Q, *Nme*F114R/R117Q and *Nme*F114R/R117Q/F139G at the Aus-

tralian Synchrotron on the Micro Crystallography (MX2) beamline, and for *Nme*F114R/R117A on the Macromolecular Crystallography (MX1) beamline.

Data were collected for crystals of *Nme*F139G grown in a 0.6 M NaCl condition, *Nme*F114R in 0.6 M to 2.8 M NaCl, *Nme*R117Q in 1.4 M NaCl, *Nme*F114R/R117Q in 0.6 M to 2.8 M NaCl, *Nme*F114R/R117Q/F139G in 2 M NaCl, and *Nme*F114R/R117A in 1.6 M NaCl. All eight mutant proteins crystallise like the wild-type *Nme*KDO8PS in orthorhombic space group $P2_12_12_1$, with the following unit cell dimensions: $a \approx 82$ Å, $b \approx 85$ Å, and $c \approx 163$ Å. The results are summarised in Tables 7.4, 7.5, 7.6, 7.7, and 7.8, along with key structure refinement details.

Appendices

Appendix A

Alignment of HMM-derived model sequence with other KDO8PS sequences

```
AfeKDO8PS   -----MRLCGFEAGLQHPFFLMAGPCAIESESLALRTAEDLRDICARLG-IPFIYKSSYD 54
EcoKDO8PS   MKQKVVSIGDINVANDLPFVLFGGMNVLESRDLAMRICEHYVTVTQKLG-IPYVFKASFD 59
AaeKDO8PS   ---------------MEKFLVIAGPCAIESEELLLKVGEEIKRLSEKFKEVEFVFKSSFD 45
NmeKDO8PS   ---MDIKINDITLGNNSPFVLFGGINVLESLDSTLQTCAHYVEVTRKLG-IPYIFKASFD 56
MostLikely  ----------------KPLVLIAGPCVLESEELALEVAEELKEITEKLG-IELVFKASFD 43
                           :.::.*  .:** .  :.   .   :  ::  :  ::*:*:*

AfeKDO8PS   KANRSSGQSFRGPGMDEGLRILEKVRREVGVPVVTDVHEKEDVSAVAEVVDVLQTPAFLC 114
EcoKDO8PS   KANRSSIHSYRGPGLEEGMKIFQELKQTFGVKIITDVHEPSQAQPVADVVDVIQLPAFLA 119
AaeKDO8PS   KANRSSIHSFRGHGLEYGVKALRKVKEEFGLKITTDIHESWQAEPVAEVADIIQIPAFLC 105
NmeKDO8PS   KANRSSIHSYRGVGLEEGLKIFEKVKAEFGIPVITDVHEPHQCQPVAEVCDVIQLPAFLA 116
MostLikely  KANRSSIKSFRGPGLEEGLKILEEVKEELGVPVLTDVHEEEQAAEVAEVVDVLQIPAFLC 103
            ****** :*:** *:: *:: :.:::  .*: : **:**  :   **:* *::* ****.

AfeKDO8PS   RQTDFIQAVAAAGKPVNIKKGQFLAPWDMLHVASKAKATGNEQIMVCERGASFGYNNLVS 174
EcoKDO8PS   RQTDLVEAMAKTGAVINVKKPQFVSPGQMGNIVDKFKEGGNEKVILCDRGANFGYDNLVV 179
AaeKDO8PS   RQTDLLLAAAKTGRAVNVKKGQFLAPWDTKNVVEKLKFGGAKEIYLTERGTTFGYNNLVV 165
NmeKDO8PS   RQTDLVVAMAKTGNVVNIKKPQFLSPSQMKNIVEKFHEAGNGKLILCERGSSFGYDNLVV 176
MostLikely  RQTDLLVAAAKTGKVVNVKKGQFLAPEDMKNVVEKVKEAGNEKVLLTERGASFGYNNLVV 163
            ****:: * * :*  :*:** **::* :  ::..* :  *  :: : :**:.***:***

AfeKDO8PS   DMRSLAVMRQTG--CPVVFDATHSVQLPG-GQGDRSGGQREFIPVLARAAVAAGVSGLFM 231
EcoKDO8PS   DMLGFSIMKKVSGNSPVIFDVTHALQCRD-PFGAASGGRRAQVAELARAGMAVGLAGLFI 238
AaeKDO8PS   DFRSLPIMKQW---AKVIYDATHSVQLPG-GLGDKSGGMREFIFPLIRAAVAVGCDGVFM 221
NmeKDO8PS   DMLGFGVMKQTCGNLPVIFDVTHSLQTRD-AGSAASGGRRAQALDLALAGMATRLAGLFL 235
MostLikely  DMRSLAIMRELG--APVIFDATHSVQLPGDGAGASSGGQREFVPELARAAVAVGVAGLFL 221
            *: .: :*::      *::*.**::*  .   .  *** *     *  *.:*.   *:*:

AfeKDO8PS   ETHPNPADALSDGPNAWPLGRMEDLLRILQHIDHVVKNQDFPENYPEELV 281
EcoKDO8PS   EAHPDPEHAKCDGPSALPLAKLEPFLKQMKAIDDLVKGFEELDTSK---- 284
AaeKDO8PS   ETHPEPEKALSDASTQLPLSQLEGIIEAILEIREVAS--KYYETIPVK-- 267
NmeKDO8PS   ESHPDPKLAKCDGPSALPLHLLEDFLIRIKALDDLIKSQPILTIE----- 280
MostLikely  ETHPDPDKAKSDGPNALPLKKLEELLEELKAIDE---------------- 255
            *:**:*  * .*...  **  :* ::  :  : .
```

Figure A.1: Alignment of the most-likely KDO8PS sequence (numbering from Chapter 2), with common KDO8PS sequences referred to in this thesis. Conservation is indicated using standard Clustal codes.

Appendix B

Mass spectrometry measurements

Table B.1: The subunit molecular weights of wild-type and mutant *Afe*KDO8PS and *Nme*KDO8PS. The theoretical molecular weights were calculated from the amino acid sequence using ProtParam on the ExPASy server.

KDO8PS	Theoretical (Da)	Measured (Da)
*Afe*WT	30 603.9	30 609.0
*Afe*C21N	30 614.9	30 613.4
*Afe*D243E	30 617.9	30 616.8
*Afe*P245A	30 577.9	30 577.4
*Afe*K55A	30 546.8	30 546.7
*Afe*N57A	30 560.9	30 561.1
*Afe*KARS	30 489.8	30 490.6
*Afe*KPRS	30 515.9	30 516.7
*Afe*Q198A	30 546.9	30 546.3
*Afe*S207A	30 587.9	30 588.5
*Afe*Q198A/S207A	30 530.9	30 530.4
*Afe*L7trun	29 550.8	29 550.9

continued on next page

Table B.1 – continued from previous page

Protein	Theoretical (Da)	Measured (Da)
*Afe*L7trun/KPRS	29 462.8	29 462.1
*Afe*D90A	30 559.9	30 559.4
*Afe*D90N	30 602.9	30 602.5
*Afe*D90E	30 617.9	30 618.3
*Nme*WT	30 483.3	30 483.0
*Nme*K57A	30 426.2	30 427.4
*Nme*N59A	30 440.3	30 441.0
*Nme*N59D	30 484.3	30 485.5
*Nme*KARS	30 369.2	30 370.9
*Nme*KPRS	30 395.2	30 394.0
*Nme*Q202A	30 426.3	30 427.9
*Nme*S211A	30 466.9	30 467.0
*Nme*Q202A/S211A	30 410.3	30 409.9
*Nme*L7trun	29 481.3	29 481.4
*Nme*L7trun/KPRS	29 393.2	29 392.4
*Nme*D92A	30 439.3	30 438.6
*Nme*D92N	30 482.3	30 482.1
*Nme*D92E	30 497.3	30 497.0
*Nme*F114A	30 407.2	30 406.8
*Nme*F114R	30 492.3	30 492.0
*Nme*F139G	30 393.2	30 392.9
*Nme*F114R/R117A	30 398.2	30 407.0
*Nme*F114R/R117Q	30 464.3	30 464.0
*Nme*F114R/R117Q/F139G	30 374.1	30 373.8
*Nme*R117Q	30 454.9	30 454.9
*Nme*KPRS/F114R/R117Q/ A126T/F139G/L7trun	29 314.1	29 313.1

References

[1] Nikaido, H. In *In Escherichia coli and Salmonella typhimurium: cellular and molecular biology.*, 2nd ed.; Neidhardt, F., Ed.; American Society for Microbiology, Washington, DC, 1996; pp 29–47.

[2] Wiese, A., Brandenburg, K., Ulmer, A. J., Seydel, U., and Müller-Loennies, S. (1999) The dual role of lipopolysaccharide as effector and target molecule. *Biol. Chem. 380*, 767–784.

[3] Heine, H., Rietschel, E., and Ulmer, A. (2001) The biology of endotoxin. *Mol. Biotechnol. 19*, 279–296.

[4] Raetz, C. R. H., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem. 71*, 635–700.

[5] Holst, O. (2002) Chemical structure of the core region of lipopolysaccharides — an update. *Trends Glycosi. Glyc. 14*, 87–103.

[6] Wyckoff, T. J. O., Raetz, C. R. H., and Jackman, J. E. (1998) Antibacterial and anti-inflammatory agents that target endotoxin. *Trends Microbiol. 6*, 154–159.

[7] Meredith, T. C., Aggarwal, P., Mamat, U., Lindner, B., and Woodard, R. W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS Chem. Biol. 1*, 33–42.

[8] Luke, N. R., Allen, S., Gibson, B. W., and Campagnari, A. A. (2003) Identification of a 3-deoxy-D-*manno*-octulosonic acid biosynthetic op-

eron in *Moraxella catarrhalis* and analysis of a KdsA-deficient isogenic mutant. *Infect. Immun. 71*, 6426–6434.

[9] Raetz, C. R. H. (1990) Biochemistry of endotoxins. *Annu. Rev. Biochem. 59*, 129–170.

[10] Gronow, S., and Brade, H. (2001) Invited review: Lipopolysaccharide biosynthesis: which steps do bacteria need to survive? *J. Endotoxin Res. 7*, 3–23.

[11] Rick, P. D., and Young, D. A. (1982) Isolation and characterization of a temperature-sensitive lethal mutant of *Salmonella typhimurium* that is conditionally defective in 3-deoxy-D-*manno*-octulosonate-8-phosphate synthesis. *J. Bacteriol. 150*, 447–455.

[12] Becker, B., Lommerse, J. P., Melkonian, M., Kamerling, J. P., and Vliegenthart, J. F. (1995) The structure of an acidic trisaccharide component from a cell wall polysaccharide preparation of the green alga *Tetraselmis striata* Butcher. *Carbohydr. Res. 267*, 313–321.

[13] York, W. S., Darvill, A. G., McNeil, M., and Albersheim, P. (1985) 3-Deoxy-D-*manno*-2-octulosonic acid (KDO) is a component of rhamnogalacturonan II, a pectic polysaccharide in the primary cell walls of plants. *Carbohydr. Res. 138*, 109–126.

[14] Delmas, F., Petit, J., Joubes, J., Seveno, M., Paccalet, T., Hernould, M., Lerouge, P., Mouras, A., and Chevalier, C. (2003) The gene expression and enzyme activity of plant 3-deoxy-D-*manno*-2-octulosonic acid-8-phosphate synthase are preferentially associated with cell division in a cell cycle-dependent manner. *Plant Physiol. 133*, 348–360.

[15] Ray, P. H. (1980) Purification and characterization of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthetase from *Escherichia coli*. *J. Bacteriol. 141*, 635–644.

[16] Kohen, A., Berkovich, R., Belakhov, V., and Baasov, T. (1993) Stereochemistry of the KDO8P synthase. An efficient synthesis of the 3-fluoro analogues of KDO8P. *Bioorg. Med. Chem. Lett. 3*, 1577–1582.

[17] Dotson, G. D., Nanjappan, P., Reily, M. D., and Woodard, R. W. (1993) Stereochemistry of 3-deoxyoctulosonate 8-phosphate synthase. *Biochemistry 32*, 12392–12397.

[18] Furdui, C. M., Sau, A. K., Yaniv, O., Belakhov, V., Woodard, R. W., Baasov, T., and Anderson, K. S. (2005) The use of (*E*)- and (*Z*)-phosphoenol-3-fluoropyruvate as mechanistic probes reveals significant differences between the active sites of KDO8P and DAHP synthases. *Biochemistry 44*, 7326–7335.

[19] Tao, P., Gatti, D., and Schlegel, H. (2009) The energy landscape of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochemistry 48*, 11706–11714.

[20] Kohen, A., Jakob, A., and Baasov, T. (1992) Mechanistic studies of 3-deoxy-D-*manno*-2-octulosonate-8-phosphate synthase from *Escherichia coli*. *Eur. J. Biochem. 208*, 443–449.

[21] Liang, P.-H., Lewis, J., Anderson, K. S., Kohen, A., D'Souza, F. W., Benenson, Y., and Baasov, T. (1998) Catalytic mechanism of Kdo8P synthase: Transient kinetic studies and evaluation of a putative reaction intermediate. *Biochemistry 37*, 16390–16399.

[22] Li, Z., Sau, A., Shen, S., Whitehouse, C., Baasov, T., and Anderson, K. (2003) A snapshot of enzyme catalysis using electrospray ionization mass spectrometry. *J. Am. Chem. Soc. 125*, 9938–9939.

[23] Roberts, A., Furdui, C., and Anderson, K. (2010) Observation of a chemically labile, noncovalent enzyme intermediate in the reaction of metal-dependent *Aquifex pyrophilus* KDO8PS by time-resolved mass spectrometry. *Rapid Commun. Mass Spectrom. 24*, 1919–1924.

[24] Kona, F., Xu, X., Martin, P., Kuzmic, P., and Gatti, D. L. (2007) Structural and mechanistic changes along an engineered path from metallo to nonmetallo 3-deoxy-D-*manno*-octulosonate 8-phosphate synthases. *Biochemistry 46*, 4532–4544.

[25] Liang, P.-H., Kohen, A., Baasov, T., and Anderson, K. S. (1997) Cata-

lytic mechanism of KDO8P synthase. Pre-steady-state kinetic analysis using rapid chemical quench flow methods. *Bioorg. Med. Chem. Lett. 7*, 2463–2468.

[26] D'Souza, F. W., Benenson, Y., and Baasov, T. (1997) Catalytic mechanism of KDO8P synthase: synthesis and evaluation of a putative reaction intermediate. *Bioorg. Med. Chem. Lett. 7*, 2457–2462.

[27] Wang, J., Duewel, H. S., Woodard, R. W., and Gatti, D. L. (2001) Structures of *Aquifex aeolicus* KDO8P synthase in complex with R5P and PEP, and with a bisubstrate inhibitor: Role of active site water in catalysis. *Biochemistry 40*, 15676–15683.

[28] Shulami, S., Furdui, C., Adir, N., Shoham, Y., Anderson, K. S., and Baasov, T. (2004) A reciprocal single mutation affects the metal requirement of 3-deoxy-D-*manno*-2-octulosonate-8-phosphate (KDO8P) synthases from *Aquifex pyrophilus* and *Escherichia coli*. *J. Biol. Chem. 279*, 45110–45120.

[29] Tao, P., Schlegel, H., and Gatti, D. (2010) Common basis for the mechanism of metallo and non-metallo KDO8P synthases. *J. Inorg. Biochem. 104*, 1267–1275.

[30] Dotson, G. D., Dua, R. K., Clemens, J. C., Wooten, E. W., and Woodard, R. W. (1995) Overproduction and one-step purification of *Escherichia coli* 3-deoxy-D-*manno*-octulosonic acid 8-phosphate synthase and oxygen transfer studies during catalysis using isotopic-shifted heteronuclear NMR. *J. Biol. Chem. 270*, 13698–13705.

[31] Hedstrom, L., and Abeles, R. (1988) 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase catalyzes the C-O bond cleavage of phosphoenolpyruvate. *Biochem. Biophys. Res. Commun. 157*, 816–820.

[32] Wooten, E. W., Dua, R. K., Dotson, G. D., and Woodard, R. W. (1994) Homo- and heteronuclear multiple-quantum filters for measurement of NMR isotope shifts. *J. Magn. Reson. 107*, 50–55.

[33] Kim, D. H., Lees, W. J., Haley, T. M., and Walsh, C. T. (1995) Kin-

etic characterization of the inactivation of UDP-GlcNAc enolpyruvyl transferase by (*Z*)-3-fluorophosphoenolpyruvate: evidence for two oxocarbenium ion intermediates in enolpyruvyl transfer catalysis. *J. Am. Chem. Soc. 117*, 1494–1502.

[34] Alberg, D. G., Lauhon, C. T., Nyfeler, R., Faessler, A., and Bartlett, P. A. (1992) Inhibition of 5-enolpyruvoylshikimate 3-phosphate (EPSP) synthase by analogs of the tetrahedral intermediate and of EPSP. *J. Am. Chem. Soc. 114*, 3535–3546.

[35] Walsh, C. T., Benson, T. E., Kim, D. H., and Lees, W. J. (1996) The versatility of phosphoenolpyruvate and its vinyl ether products in biosynthesis. *Chem. Biol. 3*, 83–91.

[36] Asojo, O., Friedman, J., Adir, N., Belakhov, V., Shoham, Y., and Baasov, T. (2001) Crystal structures of KDO8P synthase in its binary complexes with the substrate phosphoenolpyruvate and with a mechanism-based inhibitor. *Biochemistry 40*, 6326–6334.

[37] Benenson, Y., Belakhov, V., and Baasov, T. (1996) 1-(Dihydroxyphosphynyl)vinyl phosphate: the phosphonate analogue of phosphoenolpyruvate is a pH-dependent substrate of Kdo8P synthase. *Bioorg. Med. Chem. Lett. 6*, 2901–2906.

[38] Du, S., Plat, D., Belakhov, V., and Baasov, T. (1997) First nonenzymatic synthesis of Kdo8P through a mechanism similar to that suggested for the enzyme Kdo8P synthase. *J. Org. Chem. 62*, 794–804.

[39] Cochrane, F., Cookson, T., Jameson, G., and Parker, E. (2009) Reversing evolution: re-establishing obligate metal ion dependence in a metal-independent KDO8P synthase. *J. Mol. Biol. 390*, 646–661.

[40] Duewel, H. S., Sheflyan, G. Y., and Woodard, R. W. (1999) Functional and biochemical characterization of a recombinant 3-deoxy-D-*manno*-octulosonic acid 8-phosphate synthase from the hyperthermophilic bacterium *Aquifex aeolicus*. *Biochem. Biophys. Res. Commun. 263*, 346–351.

[41] Krosky, D. J., Alm, R., Berg, M., Carmel, G., Tummino, P. J., Xu, B., and Yang, W. (2002) *Helicobacter pylori* 3-deoxy-D-*manno*-octulosonate-8-phosphate (KDO-8-P) synthase is a zinc-metalloenzyme. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 1594*, 297–306.

[42] Wu, J., Patel, M. A., Sundaram, A. K., and Woodard, R. W. (2004) Functional and biochemical characterization of a recombinant *Arabidopsis thaliana* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochem. J. 381*, 185–193.

[43] Subramaniam, P. S., Xie, G., Xia, T., and Jensen, R. A. (1998) Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases. *J. Bacteriol. 180*, 119–127.

[44] Sheflyan, G. Y., Sundaram, A. K., Taylor, W. P., and Woodard, R. W. (2000) Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* revisited. *J. Bacteriol. 182*, 5005–5008.

[45] Birck, M. R., and Woodard, R. W. (2001) *Aquifex aeolicus* 3-deoxy-D-*manno*-2-octulosonic acid 8-phosphate synthase: a new class of KDO 8-P synthase? *J. Mol. Evol. 52*, 205–214.

[46] Taylor, W. P., Sheflyan, G. Y., and Woodard, R. W. (2000) A single point mutation in 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase is responsible for temperature sensitivity in a mutant strain of *Salmonella typhimurium*. *J. Biol. Chem. 275*, 32141–32146.

[47] Allison, T. M., Yeoman, J. A., Hutton, R. D., Cochrane, F. C., Jameson, G. B., and Parker, E. J. (2010) Specificity and mutational analysis of the metal-dependent 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Acidithiobacillus ferrooxidans*. *Biochim. Biophys. Acta, Proteins Proteomics 1804*, 1526–1536.

[48] Bentley, R., and Haslam, E. (1990) The shikimate pathway—a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol. 25*, 307–384.

[49] Herrmann, K. M., and Weaver, L. M. (1999) The shikimate pathway. *Annu. Rev. Plant. Physiol. Plant Mol. Biol. 50*, 473–503.

[50] Wu, J., and Woodard, R. W. (2006) New insights into the evolutionary links relating to the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase subfamilies. *J. Biol. Chem. 281*, 4042–4048.

[51] Tribe, D. E., Camakaris, H., and Pittard, J. (1976) Constitutive and repressive enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. *J. Bacteriol. 127*, 1085–1097.

[52] Hoffmann, P. J., Doy, C. H., and Catcheside, D. E. A. (1972) The separation of three allosterically inhibitable 3-deoxy-*arabino*-heptulosonate 7-phosphate synthases from extracts of *Neurospora crassa* and the purification of the tyrosine inhibitable isoenzyme. *Biochim. Biophys. Acta, Enzymol. 268*, 550–561.

[53] Hartmann, M., Schneider, T., Pfeil, A., Heinrich, G., Lipscomb, W., and Braus, G. (2003) Evolution of feedback-inhibited β/α barrel isoenzymes by gene duplication and a single mutation. *Proc. Natl. Acad. Sci. U. S. A. 100*, 862–867.

[54] Webby, C., Jiao, W., Hutton, R., Blackmore, N., Baker, H., Baker, E., Jameson, G., and Parker, E. (2010) Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem. 285*, 30567–30576.

[55] Shumilin, I. A., Zhao, C., Bauerle, R., and Kretsinger, R. H. (2002) Allosteric inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J. Mol. Biol. 320*, 1147–1156.

[56] Webby, C., Patchett, M., and Parker, E. (2005) Characterization of

a recombinant type II 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Helicobacter pylori*. *Biochem. J. 390*, 223.

[57] Wu, J., Howe, D., and Woodard, R. (2003) *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase. *J. Biol. Chem. 278*, 27525.

[58] Cross, P. J., Dobson, R. C. J., Patchett, M. L., and Parker, E. J. (2011) Tyrosine-latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J. Biol. Chem. 286*, 10216–10224.

[59] Schofield, L. R., Anderson, B. F., Patchett, M. L., Norris, G. E., Jameson, G. B., and Parker, E. J. (2005) Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: an ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry 44*, 11950–11962.

[60] Onderka, D. K., and Floss, H. G. (1969) Steric course of the chorismate synthetase reaction and the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthetase reaction. *J. Am. Chem. Soc. 91*, 5894–5896.

[61] Floss, H. G., Onderka, D. K., and Carroll, M. (1972) Stereochemistry of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase reaction and the chorismate synthetase reaction. *J. Biol. Chem. 247*, 736–744.

[62] Staub, M., and Dénes, G. (1967) A kinetic study of the mechanism of action of 3-deoxy-*arabino*-heptulosonate 7-phosphate synthase in *Escherichia coli* K 12. *Biochim. Biophys. Acta, Enzymol. 132*, 528–530.

[63] DeLeo, A. B., Dayan, J., and Sprinson, D. B. (1973) Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*. *J. Biol. Chem. 248*, 2344–2353.

[64] Jensen, R., Xie, G., Calhoun, D., and Bonner, C. (2002) The correct phylogenetic relationship of KdsA (3-deoxy-D-*manno*-octulosonate 8-phosphate synthase) with one of two independently evolved classes of

AroA (3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase). *J. Mol. Evol. 54*, 416–423.

[65] Shulami, S., Yaniv, O., Rabkin, E., Shoham, Y., and Baasov, T. (2003) Cloning, expression, and biochemical characterization of 3-deoxy-D-*manno*-2-octulosonate-8-phosphate (KDO8P) synthase from the hyperthermophilic bacterium *Aquifex pyrophilus*. *Extremophiles 7*, 471–481.

[66] Sau, A., Li, Z., and Anderson, K. (2004) Probing the role of metal ions in the catalysis of *Helicobacter pylori* 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase using a transient kinetic analysis. *J. Biol. Chem. 279*, 15787–15794.

[67] Li, J., Wu, J., Fleischhacker, A. S., and Woodard, R. W. (2004) Conversion of *Aquifex aeolicus* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase, a metalloenzyme, into a nonmetalloenzyme. *J. Am. Chem. Soc. 126*, 7448–7449.

[68] Oliynyk, Z., Briseno-Roa, L., Janowitz, T., Sondergeld, P., and Fersht, A. (2004) Designing a metal-binding site in the scaffold of *Escherichia coli* KDO8PS. *Protein Eng., Des. Sel. 17*, 383–390.

[69] Ahn, M., Pietersma, A. L., Schofield, L. R., and Parker, E. J. (2005) Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions. *Org. Biomol. Chem. 3*, 4046–4049.

[70] Stephens, C. M., and Bauerle, R. (1991) Analysis of the metal requirement of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *J. Biol. Chem. 266*, 20810–20817.

[71] Stephens, C. M., and Bauerle, R. (1992) Essential cysteines in 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. Analysis by chemical modification and site-directed mutagenesis of the phenylalanine-sensitive isozyme. *J. Biol. Chem. 267*, 5762–5767.

[72] Duewel, H. S., Radaev, S., Wang, J., Woodard, R. W., and Gatti, D. L. (2001) Substrate and metal complexes of 3-deoxy-D-*manno*-octulosonate-

8-phosphate synthase from *Aquifex aeolicus* at 1.9-A resolution. Implications for the condensation mechanism. *J. Biol. Chem. 276*, 8393–8402.

[73] Kona, F., Tao, P., Martin, P., Xu, X., and Gatti, D. (2009) Electronic structure of the metal center in the Cd^{2+}, Zn^{2+}, and Cu^{2+} substituted forms of KDO8P synthase: implications for catalysis. *Biochemistry 48*, 3610–3630.

[74] Shumilin, I. A., Bauerle, R., Wu, J., Woodard, R. W., and Kretsinger, R. H. (2004) Crystal structure of the reaction complex of 3-deoxy-*arabino*-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *J. Mol. Biol. 341*, 455–466.

[75] Reichau, S., Jiao, W., Walker, S. R., Hutton, R. D., Baker, E. N., and Parker, E. J. (2011) Potent inhibitors of a shikimate pathway enzyme from *Mycobacterium tuberculosis*. *J. Biol. Chem. 286*, 16197–16207.

[76] König, V., Pfeil, A., Braus, G. H., and Schneider, T. R. (2004) Substrate and metal complexes of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J. Mol. Biol. 337*, 675–690.

[77] Shumilin, I. A., Bauerle, R., and Kretsinger, R. H. (2003) The high-resolution structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate. *Biochemistry 42*, 3766–3776.

[78] Wang, J., Duewel, H. S., Stuckey, J. A., Woodard, R. W., and Gatti, D. L. (2002) Function of His185 in *Aquifex aeolicus* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *J. Mol. Biol. 324*, 205–214.

[79] Xu, X., Kona, F., Wang, J., Lu, J., Stemmler, T., and Gatti, D. L. (2005) The catalytic and conformational cycle of *Aquifex aeolicus* KDO8P synthase: role of the L7 loop. *Biochemistry 44*, 12434–12444.

[80] Howe, D. L., Sundaram, A. K., Wu, J., Gatti, D. L., and Woodard, R. W. (2003) Mechanistic insight into 3-deoxy-D-*manno*-octulosonate-8-

phosphate synthase and 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase utilizing phosphorylated monosaccharide analogues. *Biochemistry 42*, 4843–4854.

[81] Ahn, M., Cochrane, F. C., Patchett, M. L., and Parker, E. J. (2008) Arabinose 5-phosphate analogues as mechanistic probes for *Neisseria meningitidis* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Bioorg. Med. Chem. 16*, 9830–9836.

[82] Duewel, H. S., and Woodard, R. W. (2000) A metal bridge between two enzyme families. 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase from *Aquifex aeolicus* requires a divalent metal for activity. *J. Biol. Chem. 275*, 22824–22831.

[83] Sheflyan, G. Y., Howe, D. L., Wilson, T. L., and Woodard, R. W. (1998) Enzymatic synthesis of 3-deoxy-D-*manno*-octulosonate 8-phosphate, 3-deoxy-D-altro-octulosonate 8-phosphate, 3,5-dideoxy-D-gluco(*manno*)-octulosonate 8-phosphate by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Am. Chem. Soc. 120*, 11027–11032.

[84] Yeoman, J. A. Biochemical characterization of metal-dependent 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Chlorobium tepidum* & *Acidithiobacillus ferrooxidans*. M.Sc. thesis, Massey University, 2007.

[85] Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., Söhngen, C., Stelzer, M., Thiele, J., and Schomburg, D. (2011) BRENDA, the enzyme information system in 2011. *Nucleic Acids Res. 39*, D670–D676.

[86] Ackerman, S. H., and Gatti, D. L. (2011) The contribution of coevolving residues to the stability of KDO8P synthase. *PLoS ONE 6*, e17459.

[87] Tran, D., Pietersma, A. L., Schofield, L. R., Rost, M., Jameson, G. B., and Parker, E. J. (2011) Investigating the role of the hydroxyl groups of substrate erythrose 4-phosphate in the reaction catalysed by the first enzyme of the shikimate pathway. *Bioorg. Med. Chem. Lett. 21*, 6838–6841.

[88] Gosset, G., Bonner, C., and Jensen, R. (2001) Microbial origin of plant-type 2-keto-3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from *Xanthomonas campestris*. *J. Bacteriol. 183*, 4061.

[89] Webby, C., Baker, H., Lott, J., Baker, E., and Parker, E. (2005) The structure of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J. Mol. Biol. 354*, 927–939.

[90] Kelly, D. P., and Wood, A. P. (2000) Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. *Int. J. Syst. Evol. Microbiol. 50*, 511–516.

[91] McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics 16*, 404–405.

[92] Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London. *Nucleic Acids Res. 33*, W36–W38.

[93] Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif. 41*, 207–234.

[94] Kaustov, L., Kababya, S., Du, S., Baasov, T., Gropper, S., Shoham, Y., and Schmidt, A. (2000) Structural and mechanistic investigation of 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase by solid-state REDOR NMR. *Biochemistry 39*, 14865–14876.

[95] Allison, T. M., Hutton, R. D., Jiao, W., Gloyne, B. J., Nimmo, E. B., Jameson, G. B., and Parker, E. J. (2011) An extended β7α7 substrate-binding loop is essential for efficient catalysis by 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochemistry 50*, 9318–9327.

[96] Nimmo, E. B. The characterisation of 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase (KDO8PS) from *Neisseria meningitidis* and *Acidi-*

thiobacillus ferrooxidans and the solubility trials of *Helicobacter pylori* KDO8PS. B.Sc. Hons. thesis, University of Canterbury, 2009.

[97] Parker, E. J. Personal communication between Timothy M. Allison and Emily J. Parker. 2011.

[98] Nagano, N., Orengo, C. A., and Thornton, J. M. (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol. 321*, 741–765.

[99] Gerlt, J. A., and Raushel, F. M. (2003) Evolution of function in $(\beta/\alpha)_8$-barrel enzymes. *Curr. Opin. Chem. Biol. 7*, 252–264.

[100] Babbitt, P. C., and Gerlt, J. A. (1997) Understanding enzyme superfamilies. *J. Biol. Chem. 272*, 30591–30594.

[101] Gerlt, J. A., and Babbitt, P. C. (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annu. Rev. Biochem. 70*, 209–246.

[102] Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry 42*, 8387–8393.

[103] Vick, J. E., Schmidt, D. M. Z., and Gerlt, J. A. (2005) Evolutionary potential of $(\beta/\alpha)_8$-barrels: in vitro enhancement of a "new" reaction in the enolase superfamily. *Biochemistry 44*, 11722–11729.

[104] Vick, J. E., and Gerlt, J. A. (2007) Evolutionary potential of $(\beta/\alpha)_8$-barrels: stepwise evolution of a "new" reaction in the enolase superfamily. *Biochemistry 46*, 14589–14597.

[105] Ma, H., and Penning, T. M. (1999) Conversion of mammalian 3α-hydroxysteroid dehydrogenase to 20α-hydroxysteroid dehydrogenase

using loop chimeras: changing specificity from androgens to progestins. *Proc. Natl. Acad. Sci. U. S. A. 96*, 11161–11166.

[106] Cheon, Y.-H., Park, H.-S., Kim, J.-H., Kim, Y., and Kim, H.-S. (2004) Manipulation of the active site loops of D-hydantoinase, a $(\beta/\alpha)_8$-barrel protein, for modulation of the substrate specificity. *Biochemistry 43*, 7413–7420.

[107] Jürgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M., and Sterner, R. (2000) Directed evolution of a $(\beta\alpha)_8$-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc. Natl. Acad. Sci. U. S. A. 97*, 9925–9930.

[108] Hsu, C.-C., Hong, Z., Wada, M., Franke, D., and Wong, C.-H. (2005) Directed evolution of D-sialic acid aldolase to L-3-deoxy-*manno*-2-octulosonic acid (L-KDO) aldolase. *Proc. Natl. Acad. Sci. U. S. A. 102*, 9122–9126.

[109] Wymer, N., Buchanan, L. V., Henderson, D., Mehta, N., Botting, C. H., Pocivavsek, L., Fierke, C. A., Toone, E. J., and Naismith, J. H. (2001) Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from *Escherichia coli*. *Structure 9*, 1–9.

[110] Li, J. Substrate specificity and metal requirements of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDOPS). Ph.D. thesis, The University of Michigan, 2008.

[111] Ran, N., Draths, K. M., and Frost, J. W. (2004) Creation of a shikimate pathway variant. *J. Am. Chem. Soc. 126*, 6856–6857.

[112] Schrödinger, LLC,

[113] Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T., and Higgins, D. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics 23*, 2947–2948.

[114] Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol. 166*, 557–580.

[115] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*, 680–685.

[116] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248–254.

[117] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook* 571–607.

[118] Schoner, R., and Herrmann, K. M. (1976) 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *J. Biol. Chem. 251*, 5440–5447.

[119] Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem. 357*, 289–298.

[120] McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) *Blu-Ice* and the *Distributed Control System*: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat. 9*, 401–406.

[121] Collaborative Computational Project, Number 4, (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760–763.

[122] Murshudov, G., Vagin, A., and Dodson, E. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240–255.

[123] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126–2132.

[124] Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 12–21.

[125] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *Phaser* crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.